Weak solutions of backward stochastic differential equations with continuous generator

Nadira BOUCHEMELLA and Paul RAYNAUD DE FITTE

December 22, 2012

Abstract

We prove the existence of a weak solution to a backward stochastic differential equation (BSDE)

$$Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s) \, ds - \int_t^T Z_s \, dW_s$$

in a finite-dimensional space, where $f(t, x, y, z)$ is affine with respect to z, and satisfies a sublinear growth condition and a continuity condition. This solution takes the form of a triplet (Y, Z, L) of processes defined on an extended probability space and satisfying

$$Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s) \, ds - \int_t^T Z_s \, dW_s - (L_T - L_t)$$

where L is a martingale with possible jumps which is orthogonal to W. The solution is constructed on an extended probability space, using Young measures on the space of trajectories. One component of this space is the Skorokhod space $\mathbb{D}$ endowed with the topology S of Jakubowski.

Keywords: Backward stochastic differential equation, weak solution, martingale solution, joint solution measure, Young measure, Skorokhod space, Jakubowski's topology S, condition UT, Meyer-Zheng, pathwise uniqueness, Yamada-Watanabe-Engelbert.

Contents

1   Introduction

Aim of the paper   Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathrm{P})$ be a complete probability space, where $(\mathcal{F}_t)_{t \geq 0}$ is the natural filtration of a standard Brownian motion $W = (W_t)_{t \in [0,T]}$ on $\mathbb{R}^m$ and $\mathcal{F} = \mathcal{F}_T$.

     In this paper, we prove the existence of a weak solution (more precisely, a solution defined on an extended probability space) to the equation

$$(1) \qquad Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s - (L_T - L_t)$$

where $f(t, x, y, z)$ is affine with respect to z, and satisfies a sublinear growth condition and a continuity condition, W is an $\mathbb{R}^m$-valued standard Brownian motion, Y and Z and L are unknown processes, Y and L take their values in $\mathbb{R}^d$, Z takes its values in the space $\mathbb{L}$ of linear mappings from $\mathbb{R}^m$ to $\mathbb{R}^d$, $\xi \in \mathrm{L}^2_{\mathbb{R}^d}$ is the terminal condition, and L is a martingale orthogonal to W, with $L_0 = 0$ and with càdlàg trajectories (i.e. right continuous trajectories with left limits at every point). The process $X = (X_t)_{0 \leq t \leq T}$ is $(\mathcal{F}_t)$-adapted and continuous with values in a separable metric space $\mathbb{M}$. This process represents the random part of the generator f and plays a very small role in our construction. The space $\mathbb{M}$ can be, for example, some space of trajectories, and X_t can be, for example, the history until time t of some process ζ, i.e. $X_t = (\zeta_{s \wedge t})_{0 \leq s \leq T}$.

Such a weak solution to (1) can be considered as a generalized weak solution to the more classical equation

$$(2) \qquad Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s) \, ds - \int_t^T Z_s \, dW_s.$$

Historical comments Existence and uniqueness of the solution (Y, Z) to a nonlinear BSDE of the form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) \, ds - \int_t^T Z_s \, dW_s$$

have been proved in the seminal paper [30] by E. Pardoux and S. Peng, in the case when the generator f is random with $f(., 0, 0) \in \mathrm{L}^2(\Omega \times [0, T])$, and $f(t, y, z)$ is Lipschitz with respect to (y, z), uniformly in the other variables. In [26], J.P. Lepeltier and J. San Martín proved in the one dimensional case the existence of a solution when f is random, continuous with respect to (y, z) and satisfies a linear growth condition $\|f(t, y, z)\| \leq C(1 + \|y\| + \|z\|)$.

Equations of the form (2), with f depending on some other process X, appear in forward-backward stochastic differential equations (FBSDE's), where X is a solution of a (forward) stochastic differential equation.

As in the case of stochastic differential equations, one might expect that BSDE's with continuous generator always admit at least a *weak solution*, that is, a solution defined on a different probability space (generally with a larger filtration than the original one). A work in this direction but for forward-backward stochastic differential equations (FBSDE's) is that of K. Bahlali, B. Mezerdi, M. N'zi and Y. Ouknine [3], where the original probability is changed using Girsanov's theorem. Let us also mention the works on weak solutions to FBSDE's by Antonelli and Ma [2], and Delarue and Guatteri [13], where the change of probability space comes from the construction of the forward component.

Weak solutions where the filtration is enlarged have been studied by R. Buckdahn, H.J. Engelbert and A. Răşcanu in [11] (see also [9, 10]), using pseudopaths and the Meyer-Zheng topology [29]. Pseudopaths were invented by Dellacherie and Meyer [14], actually they are Young measures on the state space (see Subsection 3.4 for the definition of Young measures). The success of Meyer-Zheng topology comes from a tightness criterion which is easily satisfied and ensures that all limits have their trajectories in the Skorokhod space $\mathbb{D}$. We use here the fact that Meyer-Zheng's criterion also yields tightness for Jakubowski's stronger topology S on $\mathbb{D}$ [21]. Note that the result of Buckdahn, Engelbert and Răşcanu [11, Theorem 4.6] is more

general than ours in the sense that f in [11] depends functionally on Y, more precisely, their generator $f(t, x, y)$ is defined on $[0, T] \times \mathbb{D} \times \mathbb{D}$. Furthermore, in [11], W is only supposed to be a càdlàg martingale. On the other hand, it is assumed in [11] that f is bounded and does not depend on Z (but possibly on the martingale W). In the present paper, f satisfies only a linear growth condition, but the main novelty (and the main difficulty) is that f depends (linearly) on Z. As our final setup is not Brownian, the process Z we construct is not directly obtained by the martingale representation theorem, but as a limit of processes $Z^{(n)}$ which are obtained from the martingale representation theorem.

The existence of the orthogonal component L in our work comes from the fact that our approximating sequence $(Z^{(n)})$ does not converge in L^2: Actually it converges to Z only in distribution in $\mathrm{L}^2_{\mathbb{L}}[0, T]$ endowed with its weak topology, thus the stochastic integrals $\int_0^t Z^{(n)} \, dW_s$ need not converge in distribution to $\int_0^t Z \, dW_s$. Let us mention here the work of Ma, Zhang and Zheng [27], on the much more intricate problem of existence and uniqueness of weak solutions (in the classical sense) for forward-backward stochastic differential equations. Among other results, they prove existence of weak solutions with different methods and hypothesis (in particular the generator is assumed to be uniformly continuous in the space of variables) which ensure that the approximating sequence $Z^{(n)}$ constructed in their paper converges in L^2 to Z.

Let us also mention the recent paper [4] on the existence of an optimal control for a FBSDE. This optimal control and the corresponding solutions are obtained by taking weak limits of minimizing controls and the corresponding strong solutions. The limit BSDE with the optimal control also contains an orthogonal martingale component similar to ours.

In the case where the Brownian filtration needs to be enlarged, weak solutions are solutions which cannot be constructed as functionals of the sole Brownian motion W. It is natural for this construction to add some randomness to W by considering Young measures on the space of trajectories of the solutions we want to construct (let us denote momentarily Γ this space), i.e. random measures $\omega \mapsto \mu_\omega$ on Γ which depend in a measurable way on the Brownian motion. The weak solution is then constructed in the extended probability space $\underline{\Omega} = \Omega \times \Gamma$ with the probability $\mu_\omega \otimes d\,\mathrm{P}(\omega)$. Young measures have been invented many times under different names for different purposes. In the case of the construction of weak solutions of SDE's with trajectories in the Skorokhod space $\mathbb{D}$, they have been (re-)invented by Pellaumail [31] under the name of rules. In the present paper, we also

construct a weak solution with the help of Young measures on a suitable space of trajectories.

Organization of the paper   In Section 2, we give the main definitions and hypothesis, in particular we discuss and compare possible definitions of weak solutions. Using the techniques of T.G. Kurtz, we also give a Yamada-Watanabe-Engelbert type result on pathwise uniqueness and existence of strong solutions.

Section 3 is devoted to the main result, that is, the construction of a weak solution: First, we construct a sequence $(Y^{(n)}, Z^{(n)})$ of strong solutions to approximating BSDEs using a Tonelli type scheme (Subsection 3.1), then we prove uniform boudedness in L^2 of these solutions (Subsection 3.2) and compactness properties in the spaces of trajectories (Subsection 3.3). Here the space of trajectories is $\mathbb{D}_{\mathbb{R}^d}([0,T]) \times \mathrm{L}_{\mathbb{L}}^2[0,T] \times \mathbb{D}_{\mathbb{R}^d}([0,T])$, where $\mathbb{D}_{\mathbb{R}^d}([0,T])$ is endowed with Jakubowski's topology S and $\mathrm{L}_{\mathbb{L}}^2[0,T]$ with its weak topology. Finally, we obtain the solution by passing to the limit of an extracted sequence, in Young measures topology (Subsection 3.4). The proof of the main result, Theorem 3.1, is completed in Subsection 3.5.

2   General setting, weak and strong solutions

2.1   Generalities, equivalent definitions of weak solutions

Notations and hypothesis   For any separable metric space $\mathbb{E}$, we denote by $\mathrm{C}_{\mathbb{E}}[0,T]$ (respectively $\mathbb{D}_{\mathbb{E}}[0,T]$) the space of continuous (resp. càdlàg) mappings on $[0,T]$ with values in $\mathbb{E}$. (The space $\mathbb{D}_{\mathbb{R}^d}[0,T]$ will sometimes be denoted for short by $\mathbb{D}$.) Similarly, for any $q \geq 1$, if $\mathbb{E}$ is a Banach space, and if $(\Sigma, \mathcal{G}, \mathrm{Q})$ is a measure space, we denote by $\mathrm{L}_{\mathbb{E}}^q(\Sigma)$ the Banach space of measurable mappings $\varphi : \Sigma \to \mathbb{E}$ such that $\|\varphi\|_{\mathrm{L}_{\mathbb{E}}^q}^q := \int_0^T \|\varphi(s)\|^q \, d\mathrm{Q}(s) < +\infty$.

The law of a random element X of a topological space $\mathbb{E}$ is denoted by $\mathcal{L}(X)$. The conditional expectation of X with respect to a σ-algebra $\mathcal{G}$, if it exists, is denoted by $\mathrm{E}^{\mathcal{G}}(X)$. The indicator function of a set A is denoted by $\mathbb{1}_A$.

In the sequel, we are given a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathrm{P})$, the filtration $(\mathcal{F}_t)$ is the filtration generated by an $\mathbb{R}^m$-valued standard Brownian motion W, augmented with the P-negligible sets, and $\mathcal{F} = \mathcal{F}_T$. We are also given an $\mathbb{R}^d$-valued random variable $\xi \in \mathrm{L}_{\mathbb{R}^d}^2(\Omega, \mathcal{F}, \mathrm{P})$ (the terminal condition). The space of linear mappings from $\mathbb{R}^m$ to $\mathbb{R}^d$ is denoted by $\mathbb{L}$.

We denote by $\mathbb{M}$ a separable metric space and by X a given $(\mathcal{F}_t)$-adapted $\mathbb{M}$-valued continuous process. Finally we are given a measurable mapping $f : [0,T] \times \mathbb{M} \times \mathbb{R}^d \times \mathbb{L} \to \mathbb{R}^d$ which satisfies the following growth and continuity conditions (H_1) and (H_2) (which will be needed only in Section 3 for the construction of a solution):

(H_1) There exists a constant $C_f \geq 0$ such that $\forall (t,x,y,z) \in [0,T] \times \mathbb{M} \times \mathbb{R}^d \times \mathbb{L}$, $\|f(t,x,y,z)\| \leq C_f(1 + \|z\|)$.

(H_2) $f(t,x,y,z)$ is continuous with respect to (x,y) and affine with respect to z.

Weak and strong solutions

Definition 2.1 A *strong solution* to (2) is an $(\mathcal{F}_t)$-adapted, $\mathbb{R}^d \times \mathbb{L}$-valued process (Y,Z) (defined on $\Omega \times [0,T]$) which satisfies

$$(3) \qquad\qquad \int_0^T \|Z_s\|^2 \, ds < \infty \ \text{P-a.e.}$$

$$(4) \qquad\qquad \int_0^T \|f(s, X_s, Y_s, Z_s)\| \, ds < \infty \ \text{P-a.e.}$$

and such that the BSDE (2) holds true.

Remark 2.2 Similarly, a strong solution to (1) should be a triplet (Y,Z,L) defined on $\Omega \times [0,T]$) satisfying (3), (4), and (1), and such that L is a càdlàg martingale orthogonal to W and $L_0 = 0$, but this notion coincides with that of a strong solution to (2), because then L would be an $(\mathcal{F}_t)$-martingale, hence $L = 0$.

Remark 2.3 The process X is given and $(\mathcal{F}_t)$-adapted, and the final condition ξ is given and $\mathcal{F}_T$-measurable. By a well known result due to Doob (see [15, page 603] or [14, page 18]), there exists thus a Borel-measurable mapping $F : C_{\mathbb{R}^m}[0,T] \to C_{\mathbb{M}}[0,T] \times \mathbb{R}^d$ such that $(X,\xi) = F(W)$ a.e. In other words, the law $\mathcal{L}(W,X,\xi)$ of (W,X,ξ) is supported by the graph of F. The fact that X is $(\mathcal{F}_t)$-adapted is a property of F: it means that, for every $t \in [0,T]$, the restriction of X to $[0,t]$ only depends on the restriction of W to $[0,t]$.

We now give three equivalent definitions of a weak solution:

Definition 2.4

1) A *weak solution* to (1) is a stochastic basis $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \leq t \leq T}, \mu)$ along with a list $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X})$ of processes defined on $\underline{\Omega} \times [0, T]$, and adapted to $(\underline{\mathcal{F}}_t)$, and a random variable $\underline{\xi}$ defined on $\underline{\Omega}$, such that:

(W1) The processes $\underline{W}$ and $\underline{X}$ are continuous with values in $\mathbb{R}^m$ and $\mathbb{M}$ respectively, $\underline{\xi}$ takes its values in $\mathbb{R}^d$, and the law of $(\underline{W}, \underline{X}, \underline{\xi})$ on $C_{\mathbb{R}^m}[0, T] \times C_{\mathbb{M}}[0, T] \times \mathbb{R}^d$ is that of (W, X, ξ).

(W2) $\underline{W}$ is a standard Brownian motion with respect to the filtration $(\underline{\mathcal{F}}_t)$.

(W3) The processes $\underline{Y}$ and $\underline{L}$ are $\mathbb{R}^d$-valued and càdlàg, and $\underline{Z}$ is $\mathbb{L}$-valued, with $E \int_0^T \|\underline{Z}_s\|^2 \, ds < \infty$, the process $\underline{L}$ is a square integrable martingale with $\underline{L}_0 = 0$, and $\underline{L}$ is orthogonal to $\underline{W}$.

(W4) Condition (4) and the BSDE (1) hold true, replacing Y, Z, L, W, X, ξ by $\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi}$.

We then say that $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ is a *weak solution defined on* $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t), \mu)$.

2) Following the terminology of [17, 16, 25], and with the preceding notations, the probability measure $\mathcal{L}\left(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi}\right)$ on $\mathbb{D}_{\mathbb{R}^d}[0, T] \times L^2_{\mathbb{L}}[0, T] \times \mathbb{D}_{\mathbb{R}^d}[0, T] \times C_{\mathbb{R}^m}[0, T] \times C_{\mathbb{M}}[0, T] \times \mathbb{R}^d$ is called a *joint solution measure to* (1) *generated by* $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$ *and* $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$. (Here the Borel subsets of $\mathbb{D}_{\mathbb{R}^d}[0, T]$ are generated by the projection mappings $\pi_t : x \mapsto x(t)$ for $t \in [0, T]$; we shall see later that these sets are the Borel sets of the topology S of A. Jakubowski [21].)

3) An *extended solution* to (1) consists of a stochastic basis $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \leq t \leq T}, \mu)$ along with a triplet (Y, Z, L) of processes defined on $\underline{\Omega}$ such that:

(E1) There exists a measurable space $(\Gamma, \mathcal{G})$, and a filtration $(\mathcal{G}_t)$ on $(\Gamma, \mathcal{G})$ such that

$$\underline{\Omega} = \Omega \times \Gamma, \quad \underline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{G}, \quad \underline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{G}_t \text{ for every } t,$$

and there exists a probability measure μ on $(\underline{\Omega}, \underline{\mathcal{F}})$ such that $\mu(A \times \Gamma) = P(A)$ for every $A \in \mathcal{F}$.

Note that every random variable ζ defined on Ω can then be identified with a random variable defined on $\underline{\Omega}$, by setting $\zeta(\omega, \gamma) = \zeta(\omega)$. Furthermore, $\mathcal{F}$ can be viewed as a sub-σ-algebra of $\underline{\mathcal{F}}$ by identifying each $A \in \mathcal{F}$ with the set $A \times \Gamma$. Similarly, each $\mathcal{F}_t$ can be considered as a sub-σ-algebra of $\underline{\mathcal{F}}_t$. We say that $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$ is an *extension* of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$.

(E2) The process $(W_t)_{0 \le t \le T}$ is a Brownian motion on $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$ (where $W(\omega, \gamma) := W(\omega)$ for all $(\omega, \gamma) \in \underline{\Omega}$),

(E3) The processes Y, Z and L are $(\underline{\mathcal{F}}_t)$-adapted, Y and L are $\mathbb{R}^d$-valued and càdlàg, and Z is $\mathbb{L}$-valued, with $\mathrm{E} \int_0^T \|Z_s\|^2 \, ds < \infty$, and L is a square integrable martingale with $L_0 = 0$, and L is orthogonal to W.

(E4) Condition (4) and the BSDE (1) hold true.

Obviously, an extended solution is a weak solution, and a weak solution generates a joint solution measure. Actually, these concepts are equivalent in the sense that:

Proposition 2.5 *Given a joint solution measure ν to (1), there exists an extended solution to (1) which generates ν.*

Before we give the proof of Proposition 2.5, let us give an intrinsic characterization of joint solution measures. Let us first observe that:

1. It is easy to check (see the proof of Lemma 3.3) that, if Y, Z, L, W, X, ξ are defined on a stochastic basis $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t), \mu)$, then (1) is equivalent to

$$(5) \qquad Y_t = \mathrm{E}^{\underline{\mathcal{F}}_t} \left(\xi + \int_t^T f(s, X_s, Y_s, Z_s) \, ds \right)$$

$$\int_0^t Z_s \, dW_s + L_t = \mathrm{E}^{\underline{\mathcal{F}}_t} \left(\xi + \int_0^T f(s, X_s, Y_s, Z_s) \, ds \right)$$

$$(6) \qquad \qquad \qquad - \mathrm{E}^{\underline{\mathcal{F}}_0} \left(\xi + \int_0^T f(s, X_s, Y_s, Z_s) \, ds \right).$$

2. If $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ is a weak solution defined on a stochastic basis $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t), \mu)$, it is still a weak solution if we reduce the filtration $(\underline{\mathcal{F}}_t)$ to a filtration $(\underline{\mathcal{F}}'_t)$ such that $\underline{\mathcal{F}}'_t \subset \underline{\mathcal{F}}_t$ for every t and $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X})$ remains adapted to $(\underline{\mathcal{F}}'_t)$. Furthermore, conditions (W1) to (W4) remain unchanged if we augment $(\underline{\mathcal{F}}'_t)$ with the μ-negligible sets. So, if ν is a joint solution measure, there exists $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ defined on a stochastic basis $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \le t \le T}, \mu)$ such that

(W0) $\nu = \mathcal{L} \left(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi} \right)$ and $(\underline{\mathcal{F}}_t) = (\mathcal{F}_t^{\underline{Y}, \underline{Z}, \underline{L}, \underline{W}})$, where $(\mathcal{F}_t^{\underline{Y}, \underline{Z}, \underline{L}, \underline{W}})$ is the filtration generated by $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W})$, augmented with the μ-negligible sets.

Now, Condition (W1) is clearly a condition on ν. Let us rewrite Conditions (W2)-(W4), under Assumption (W0) on $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ and $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \leq t \leq T}, \mu)$. We use here techniques of Kurtz [25].

• By Lévy's characterization of Brownian motion, (W2) is satisfied if and only if $\underline{W}$ is an $(\underline{\mathcal{F}}_t)$-martingale and $\left[\underline{W}^{[i]}, \underline{W}^{[j]}\right]_t = \delta_{ij} t$ for all $t \in [0, T]$, where δ_{ij} is the Kronecker symbol and $\underline{W}^{[i]}$ is the ith coordinate of $\underline{W}$. The latter condition is satisfied if $\underline{W}$ has the same law as W, thus it is implied by (W1). We can thus replace (W2) by

(W2') $\underline{W}$ is an $(\mathcal{F}_t^{\underline{Y}, \underline{Z}, \underline{L}, \underline{W}})$-martingale.

But (W2') is equivalent to

$$(7) \qquad \mathrm{E}\big((\underline{W}_T - \underline{W}_t)\, h(\underline{Y}_{.\wedge t}, \underline{Z}_{.\wedge t}, \underline{L}_{.\wedge t}, \underline{W}_{.\wedge t})\big) = 0$$

for every $t \in [0, T]$ and every bounded Borel measurable function h defined on $\mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{L}^2_{\mathbb{L}}[0, T] \times \mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{C}_{\mathbb{R}^m}[0, T]$.

• By (W0), $\underline{Y}$ and $\underline{L}$ have their trajectories in $\mathbb{D}_{\mathbb{R}^d}[0, T]$ and $\underline{Z}$ in $\mathrm{L}^2_{\mathbb{L}}[0, T]$, and we have $\underline{L}_0 = 0$. Thus we can replace (W3) by

(W3') $\underline{L}$ is a square integrable $(\mathcal{F}_t^{\underline{Y}, \underline{Z}, \underline{L}, \underline{W}})$-martingale and $\underline{L}$ is orthogonal to $\underline{W}$.

The second part of (W3') means that $\underline{L}^{(i)} \underline{W}^{(j)}$ is a martingale for every $i \in \{1, \ldots, d\}$ and every $j \in \{1, \ldots, m\}$, where $\underline{L}^{(i)}$ and $\underline{W}^{(j)}$ denote the coordinate processes. Thus (W3') can be expressed as

$$(8) \qquad \mathrm{E}\bigg(\big(\underline{L}_T - \underline{L}_t\big) h(\underline{Y}_{.\wedge t}, \underline{Z}_{.\wedge t}, \underline{L}_{.\wedge t}, \underline{W}_{.\wedge t}) \bigg) = 0$$

$$(9) \qquad \mathrm{E}\bigg(\Big(\underline{L}_T^{(i)} \underline{W}_T^{(j)} - \underline{L}_t^{(i)} \underline{W}_t^{(j)}\Big) h(\underline{Y}_{.\wedge t}, \underline{Z}_{.\wedge t}, \underline{L}_{.\wedge t}, \underline{W}_{.\wedge t}) \bigg) = 0$$

for every $t \in [0, T]$ and every bounded Borel measurable function h defined on $\mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{L}^2_{\mathbb{L}}[0, T] \times \mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{C}_{\mathbb{R}^m}[0, T]$.

• Under (W0), Equations (5) and (6) amount to

$$\underline{Y}_t = \mathrm{E}^{\mathcal{F}_t^{(\underline{Y}, \underline{Z}, \underline{L}, \underline{W})}} \left(\underline{\xi} + \int_t^T f(s, \underline{X}_s, \underline{Y}_s, \underline{Z}_s)\, ds \right)$$

$$\int_0^t \underline{Z}_s\, d\underline{W}_s + \underline{L}_t = \mathrm{E}^{\mathcal{F}_t^{(\underline{Y}, \underline{Z}, \underline{L}, \underline{W})}} \left(\underline{\xi} + \int_t^T f(s, \underline{X}_s, \underline{Y}_s, \underline{Z}_s)\, ds \right)$$

$$- \mathrm{E}^{\mathcal{F}_0^{(\underline{Y}, \underline{Z}, \underline{L}, \underline{W})}} \left(\underline{\xi} + \int_t^T f(s, \underline{X}_s, \underline{Y}_s, \underline{Z}_s)\, ds \right).$$

9

Thus Condition (W4) is equivalent to

$$(10) \qquad \int_0^T \|f(s, \underline{X}_s, \underline{Y}_s, \underline{Z}_s)\| \, ds < \infty \ \text{P-a.e.}$$

$$(11) \ \ \mathrm{E}\left[\left(\underline{Y}_t - \underline{\xi} - \int_t^T f(s, \underline{X}_s, \underline{Y}_s, \underline{Z}_s) \, ds\right) h(\underline{Y}_{.\wedge t}, \underline{Z}_{.\wedge t}, \underline{L}_{.\wedge t}, \underline{W}_{.\wedge t})\right] = 0$$

$$(12) \quad \mathrm{E}\left[\left(\int_0^t \underline{Z}_s \, d\underline{W}_s + \underline{L}_t \right.\right.$$
$$\left. - \underline{\xi} - \int_0^T f(s, \underline{X}_s, \underline{Y}_s, \underline{Z}_s) \, ds + \mathrm{E}\left(\underline{\xi} + \int_0^T f(s, \underline{X}_s, \underline{Y}_s, \underline{Z}_s) \, ds\right)\right)$$
$$\left. \times h(\underline{Y}_{.\wedge t}, \underline{Z}_{.\wedge t}, \underline{L}_{.\wedge t}, \underline{W}_{.\wedge t})\right] = 0$$

for every bounded measurable function h defined on $\mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{L}^2_{\mathbb{L}}[0, T] \times \mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{C}_{\mathbb{R}^m}[0, T]$.

Clearly, Equations (7), (8), (9), (10), (11), and (12) only depend on the probability measure $\nu = \mathcal{L}\left(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi}\right)$. We have thus proved the following lemma, which is actually a characterization of joint solution measures:

Lemma 2.6 *Let $\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}$ be stochastic processes defined on a stochastic basis $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \le t \le T}, \mu)$, with trajectories respectively in $\mathbb{D}_{\mathbb{R}^d}[0, T]$, $\mathrm{L}^2_{\mathbb{L}}[0, T]$, $\mathbb{D}_{\mathbb{R}^d}[0, T]$, $\mathrm{C}_{\mathbb{R}^m}[0, T]$, and $\mathrm{C}_{\mathbb{M}}[0, T]$, and let $\underline{\xi}$ be an $\mathbb{R}^d$-valued random variable defined on $\underline{\Omega}$. Assume that $(\underline{\mathcal{F}}_t)$ is the filtration generated by $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W})$, possibly augmented with μ-negligible sets. Then $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ is a weak solution to (1) defined on $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \le t \le T}, \mu)$ if and only if (W1) and Equations (7), (8), (9), (10), (11), and (12) are satisfied.*

Corollary 2.7 *Let ν be a joint solution measure to (1). Let $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ and $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \le t \le T}, \mu)$ as in Lemma 2.6. Assume that $\mathcal{L}\left(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi}\right) = \nu$. Then $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ is a weak solution to (1) defined on $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{0 \le t \le T}, \mu)$.*

In particular, if ν is a joint solution measure to (1), the canonical process on the space $\mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{L}^2_{\mathbb{L}}[0, T] \times \mathbb{D}_{\mathbb{R}^d}[0, T] \times \mathrm{C}_{\mathbb{R}^m}[0, T] \times \mathrm{C}_{\mathbb{M}}[0, T] \times \mathbb{R}^d$ endowed with the probability ν is a weak solution to (1).

Before we give the proof of Proposition 2.5, let us give a definition which will be used several times. Let μ be a probability measure on a product

10

$(\Omega \times \Gamma, \mathcal{F} \otimes \mathcal{G})$ of measurable spaces such that Γ is a Polish space (or more generally, a Radon space) and $\mathcal{G}$ is its Borel σ-algebra. Let P denote the marginal measure of μ on Ω, that is, $P(A) = \mu(A \times \Gamma)$ for all $A \in \mathcal{F}$. Then there exists a unique (up to equality P-a.e.) family $(\mu_\omega)_{\omega \in \Omega}$ such that $\omega \mapsto \mu_\omega(B)$ is measurable for every $B \in \mathcal{G}$, and

$$(13) \qquad \mu(\varphi) = \int_\Omega \mu_\omega(\varphi(\omega, .)) \, d\,P(\omega)$$

for every $\mathcal{F} \otimes \mathcal{G}$-measurable nonnegative function $\varphi : \Omega \times \Gamma \to \mathbb{R}$, see e.g. [36].

Definition 2.8 The family (μ_ω) in (13) is called the *disintegration* of μ with respect to P. It is convenient to denote

$$\mu = \mu_\omega \otimes d\,P(\omega).$$

Proof of Proposition 2.5 Let $\Gamma = \mathbb{D}_{\mathbb{R}^d}[0,T] \times L^2_{\mathbb{L}}[0,T] \times \mathbb{D}_{\mathbb{R}^d}[0,T]$. Let $\mathcal{G}$ be the Borel σ-algebra of Γ, and, for each $t \in [0,T]$, let $\mathcal{G}_t$ be the σ-algebra generated by the projection of Γ onto $\mathbb{D}_{\mathbb{R}^d}[0,t] \times L^2_{\mathbb{L}}[0,t] \times \mathbb{D}_{\mathbb{R}^d}[0,t]$. Let $F : C_{\mathbb{R}^m}[0,T] \mapsto C_{\mathbb{M}}[0,T] \times \mathbb{R}^d$ be as in Remark 2.3. Then, with slight abuses of notations, ν is the image of a probability measure λ on $\Gamma \times C_{\mathbb{R}^m}[0,T] \times C_{\mathbb{M}}[0,T] \times \mathbb{R}^d$ by the mapping

$$\begin{cases} \Gamma \times C_{\mathbb{R}^m}[0,T] & \to & \Gamma \times C_{\mathbb{R}^m}[0,T] \times C_{\mathbb{M}}[0,T] \times \mathbb{R}^d \\ (y, z, l, w) & \mapsto & (y, z, l, w, F(w)). \end{cases}$$

Let $(\lambda_w)_{w \in C_{\mathbb{R}^m}[0,T]}$ be the disintegration of λ with respect to $\mathcal{L}(W)$, that is, $(\lambda_w)_{w \in C_{\mathbb{R}^m}[0,T]}$ is a family of probability measures on $(\Gamma, \mathcal{G})$ such that, for every bounded measurable $\varphi : \Gamma \times C_{\mathbb{R}^m}[0,T] \to \mathbb{R}$,

$$\lambda(\varphi) = \int_{C_{\mathbb{R}^m}[0,T]} \left(\int_\Gamma \varphi(y, z, l, w) \, d\lambda_w(y, z, l) \right) d\mathcal{L}(W)(w).$$

Now, let

$$\underline{\Omega} = \Omega \times \Gamma, \quad \underline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{G}, \quad \underline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{G}_t \quad (t \in [0,T]),$$

and let $\mu = \lambda_{W(\omega)} \otimes d\,P(\omega)$, i.e. μ is the probability measure on $(\underline{\Omega}, \underline{\mathcal{F}})$ such that

$$\mu(\varphi) = \int_\Omega \left(\int_\Gamma \varphi(\omega, y, z, l) \, d\lambda_{W(\omega)}(y, z, l) \right) d\,P(\omega)$$

11

for every bounded measurable $\varphi : \underline{\Omega} \to \mathbb{R}$. We define the random variables Y, Z, L, W, X and ξ on $\underline{\Omega}$ by

$$Y(\omega, y, z, l) = y, \quad Z(\omega, y, z, l) = z, \quad L(\omega, y, z, l) = l,$$
$$W(\omega, y, z, l) = W(\omega), \quad X(\omega, y, z, l) = X(\omega), \quad \xi(\omega, y, z, l) = \xi(\omega).$$

Then $\underline{\mathcal{F}}_t = \mathcal{F}_t^{(Y,Z,L,W)}$ for every $t \in [0, T]$, where $(\mathcal{F}_t^{Y,Z,L,W})$ is the filtration generated by (Y, Z, L, W) augmented with the P-negligible sets. Furthermore, we have $\mathcal{L}(Y, Z, L, W, X, \xi) = \nu$, thus, by Corollary 2.7, as ν is a joint solution measure to (1), (Y, Z, L, W, X, ξ) is a weak solution to (1) defined on $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t), \mu)$. □

Remark 2.9 The extension which generates ν in Proposition 2.5 is not unique. The one we construct in Section 3 is based on a different construction of the auxiliary space Γ.

We now give a criterion for an extended probability space to preserve martingales. The equivalence (ii)⇔(iii) in the following lemma is contained in Lemma 2.17 of [19].

Lemma 2.10 *Let* $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t), \mu) = (\Omega \times \Gamma, \mathcal{F} \otimes \mathcal{G}, (\mathcal{F}_t \otimes \mathcal{G}_t), \mu)$ *be an extension of* $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathrm{P})$. *Let* (μ_ω) *be the disintegration of* μ *with respect to* P. *The following are equivalent:*

(i) *W is an $(\underline{\mathcal{F}}_t)$-Brownian motion under μ,*

(ii) *Every $(\mathcal{F}_t)$-martingale is an $(\underline{\mathcal{F}}_t)$-martingale under μ[1],*

(iii) *For every $t \in [0, T]$ and every $B \in \mathcal{G}_t$, the mapping $\omega \mapsto \mu_\omega(B)$ is $\mathcal{F}_t$-measurable.*

Proof Assume (i). Let M be an $(\mathcal{F}_t)$-martingale with values in $\mathbb{R}^k$ for some integer k. Assume first that M is square integrable. By the martingale representation theorem, there exists an $(\mathcal{F}_t)$-adapted process H with $\mathrm{E} \int_0^T H_s^2 \, ds < +\infty$ such that $M_t = M_0 + \int_0^t H_s \, dW_s$. By (i), M is an $(\underline{\mathcal{F}}_t)$-martingale. In the general case, denote, for every integer $N \geq 1$,

$$M_T^N = \begin{cases} \frac{M_T}{\|M_T\|} & \text{if } \|M_T\| > N \\ M_T & \text{if } \|M_T\| \leq N, \end{cases}$$

---

[1]According to Jacod and Mémin's terminology [19, Definition 1.7], this means that $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$ is a *very good extension* of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathrm{P})$. A similar condition is called *compatibility* in [25].

and set $M_t^N = \mathrm{E}^{\mathcal{F}_t}\left(M_T^N\right)$ for $0 \le t \le T$. Then, for any $A \in \underline{\mathcal{F}}_t$, using Lebesgue's dominated convergence theorem, we have

$$\mathrm{E}\left(\mathbf{1}_A\left(M_T - M_t\right)\right) = \lim_{N\to\infty} \mathrm{E}\left(\mathbf{1}_A\left(M_T^N - M_t^N\right)\right) = 0$$

which proves that $M_t = \mathrm{E}^{\underline{\mathcal{F}}_t}(M_T)$. Thus (ii) is satisfied.

Assume (ii), and let $B \in \mathcal{G}_t$. For $u : \Omega \to \mathbb{R}$ and $v : \Gamma \to \mathbb{R}$ we denote by $u \otimes v$ the function defined on $\Omega \times \Gamma$ by $u \otimes v(\omega, x) = u(\omega)v(x)$. For each bounded $\mathcal{F}$-measurable random variable K, we have

$$\mathrm{E}\left(K\,\mu_.(B)\right) = \mu\left(K \otimes \mathbf{1}_B\right) = \mu\left(\mathrm{E}^{\underline{\mathcal{F}}_t}\left(K \otimes \mathbf{1}_B\right)\right) = \mu\left(\mathrm{E}^{\mathcal{F}_t}\left(K\right) \otimes \mathbf{1}_B\right)$$

$$(14) \qquad\qquad = \mathrm{E}\left(\mathrm{E}^{\mathcal{F}_t}\left(K\right)\mu_.(B)\right).$$

Assume that $\mu_.(B)$ is not $\mathcal{F}_t$-measurable, then, as $\mathcal{F}_t$ is complete, we have

$$0 < \mathrm{E}\left(\left(\mu_.(B) - \mathrm{E}^{\mathcal{F}_t}\left(\mu_.(B)\right)\right)^2\right) = \mathrm{E}\left(\mu_.(B)\left(\mu_.(B) - \mathrm{E}^{\mathcal{F}_t}\left(\mu_.(B)\right)\right)\right).$$

Let $K = \mu_.(B) - \mathrm{E}^{\mathcal{F}_t}\left(\mu_.(B)\right)$. We thus have

$$\mathrm{E}\left(K\mu_.(B)\right) > 0.$$

As we have, obviously,

$$\mathrm{E}\left(\mathrm{E}^{\mathcal{F}_t}\left(K\right)\mu_.(B)\right) = 0,$$

this contradicts (14). Thus $\mu_.(B)$ is $\mathcal{F}_t$-measurable, which proves (iii).

Assume (iii). To prove (i), we only need to check that W has independent increments under μ. Let $t \in [0, T]$, and let $s > 0$ such that $t + s \in [0, T]$. Let us prove that, for any $A \in \underline{\mathcal{F}}_t$ and any Borel subset C of $\mathbb{R}^m$, we have

$$(15) \qquad \mu\left(A \cap \{W_{t+s} - W_t \in C\}\right) = \mu(A)\,\mu\{W_{t+s} - W_t \in C\}.$$

Let $B = \{\omega \in \Omega;\ W_{t+s}(\omega) - W_t(\omega) \in C\}$. We have

$$\mu\left(A \cap (B \times \Gamma)\right) = \int_{\Omega \times \Gamma} \mathbf{1}_A(\omega, \gamma))\,\mathbf{1}_B(\omega)\,d\mu(\omega, \gamma)$$

$$= \int_\Omega \mu_\omega(\mathbf{1}_A(\omega, .))\,\mathbf{1}_B(\omega)\,d\,\mathrm{P}(\omega)$$

$$= \int_\Omega \mu_\omega(\mathbf{1}_A(\omega, .))\,d\,\mathrm{P}(\omega)\,\mathrm{P}(B)$$

$$= \mu(A)\,\mu(B \times \Gamma),$$

which proves (15). Thus $W_{t+s} - W_t$ is independent of $\underline{\mathcal{F}}_t$. $\qquad\square$

2.2 Pathwise uniqueness and strong solutions

One easily sees that, under hypothesis (H_1) and (H_2), Equation (2) may have infinitely many *strong* solutions. For example, let $d = m = 1$, $\xi = 0$, and $f(s,x,y,z) = \sqrt{|y|}$. Then, for any $t_0 \in [0,T]$, we get a solution by setting $Z = 0$ and

$$Y_t = \begin{cases} \frac{1}{4}(t_0 - t)^2 & \text{if } 0 \le t \le t_0 \\ 0 & \text{if } t_0 \le t \le T. \end{cases}$$

Following the usual terminology, let us say that *pathwise uniqueness* holds for Equation (1) if two weak solutions defined on the same probability space, and with respect to the same (W, X, ξ), necessarily coincide. Thus, in our setting, pathwise uniqueness does not necessarily hold.

T. G. Kurtz [25] has proved a very general version of the Yamada-Watanabe and Engelbert theorems on uniqueness and existence of strong solutions to stochastic equations, which includes SDEs, BSDEs and FBS-DEs, but without z in the generator. His results are based on the convexity of the set of joint solution-measures when the trajectories lie in a Polish space.

We can consider here that $\mathbb{D}_{\mathbb{R}^d}[0,T]$ is equipped with Skorokhod's topology J_1, which is Polish (actually, in Section 3, we will use Jakubowski's topology S on $\mathbb{D}_{\mathbb{R}^d}[0,T]$, which is not Polish, but this topology has the same Borel subsets as J_1). Thus the space $\Gamma = \mathbb{D}_{\mathbb{R}^d}[0,T] \times \mathrm{L}^2_{\mathbb{L}}[0,T] \times \mathbb{D}_{\mathbb{R}^d}[0,T]$ is Polish. In particular, Theorem 3.15 of [25] applies to our framework.

Proposition 2.11 (Yamada-Watanabe-Engelbert à la Kurtz) *Assume that pathwise uniqueness holds for Equation (1). Then every weak solution to (1) is a strong solution. Conversely, if every solution to (2) is strong, (equivalently, by Remark 2.2, if every solution to (1) is strong), then pathwise uniqueness holds for Equation (1).*

Proof In order to apply [25, Theorem 3.15], we only need to check that the set of joint solution measures to (1) is convex. (Theorem 3.15 in [25] supposes that $\mu \in S_{\Gamma,C,\nu}$ in the notations of [25], but a joint solution measure is exactly an element of $S_{\Gamma,C,\nu}$.) We check this convexity by an adaptation of [25, Example 3.17].

The set $\mathcal{M}$ of laws of joint solution measures to (1) is the set of probability laws of random variables $(\underline{Y}, \underline{Z}, \underline{L}, \underline{W}, \underline{X}, \underline{\xi})$ with values in $\mathbb{D}_{\mathbb{R}^d}[0,T] \times \mathrm{L}^2_{\mathbb{L}}[0,T] \times \mathbb{D}_{\mathbb{R}^d}[0,T] \times \mathrm{C}_{\mathbb{R}^m}[0,T] \times \mathrm{C}_{\mathbb{M}}[0,T] \times \mathbb{R}^d$, satisfying the conditions of Lemma 2.6. But each of these conditions is a convex constraint on $\mathcal{M}$.

For example, to show that Equation (12) is a convex constraint on $\mathcal{M}$, let us prove that the map $\mathcal{L}(Z, M) \mapsto \mathcal{L}\left(\int_0^\cdot Z_s \, dM_s\right)$ preserves convex combinations of probability laws. More precisely, let $\mathcal{M}^{1,+}(\mathfrak{X})$ denote the set of all probability laws on a measurable space $\mathfrak{X}$. Let $\mathcal{C}$ be the subset of $\mathcal{M}^{1,+}\left(\mathrm{L}^2_{\mathbb{L}}[0, T] \times \mathrm{C}_{\mathbb{R}^m}[0, T]\right)$ consisting of laws of processes (Z, M) such that M is a standard $\mathbb{R}^m$-valued Brownian motion and Z is $\mathbb{L}$-valued and M-adapted. We show that the mapping

$$
\begin{cases}
\mathcal{C} & \to \quad \mathcal{M}^{1,+}(\mathrm{C}_{\mathbb{R}^d}[0, T]) \\
\mathcal{L}(Z, M) & \mapsto \quad \mathcal{L}\left(\int_0^\cdot Z_s \, dM_s\right)
\end{cases}
$$

preserves convex combinations of probability laws. Indeed, Let $\mu_1, \mu_2 \in \mathcal{C}$, and let $p \in [0, 1]$. Let (Z^1, M^1) and (Z^2, M^2) be adapted processes defined on stochastic bases $(\Omega_1, \mathcal{F}^1, (\mathcal{F}^1_t), \mathrm{P}_1)$ and $(\Omega_2, \mathcal{F}^2, (\mathcal{F}^2_t), \mathrm{P}_2)$ with laws μ_1 and μ_2 respectively, such that M^1 (respectively M^2) is an $(\mathcal{F}^1_t)$-Brownian motion (resp. $(\mathcal{F}^2_t)$-Brownian motion). Let A be a random variable taking the values 1 with probability p and -1 with probability $1 - p$, defined on a probability space $(\Omega_0, \mathcal{F}^0, \mathrm{P}_0)$. We define a stochastic basis $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t), \widetilde{\mathrm{P}})$ by

$$
\widetilde{\Omega} = \Omega_0 \times \Omega_1 \times \Omega_2, \quad \widetilde{\mathcal{F}} = \mathcal{F}^0 \otimes \mathcal{F}^1 \otimes \mathcal{F}^2, \quad \widetilde{\mathcal{F}}_t = \mathcal{F}^0 \otimes \mathcal{F}^1_t \otimes \mathcal{F}^2_t,
$$
$$
\widetilde{\mathrm{P}} = \mathrm{P}_0 \otimes \mathrm{P}_1 \otimes \mathrm{P}_2.
$$

For $(\omega_0, \omega_1, \omega_2) \in \widetilde{\Omega}$, set

$$
(Z, M)(\omega_0, \omega_1, \omega_2) = \begin{cases}
(Z^1, M^1)(\omega_1) & \text{if } A(\omega_0) = 1 \\
(Z^2, M^2)(\omega_2) & \text{if } A(\omega_0) = -1.
\end{cases}
$$

Then $\mathcal{L}(Z, M) = p\mu_1 + (1 - p)\mu_2 \in \mathcal{C}$, the process M is $(\widetilde{\mathcal{F}}_t)$-Brownian, and

$$
\int_0^\cdot Z_s \, dM_s = \mathbb{1}_{\{A=1\}} \int_0^\cdot Z^1_s \, dM^1_s + \mathbb{1}_{\{A=-1\}} \int_0^\cdot Z^2_s \, dM^2_s,
$$

thus

$$
\mathcal{L}\left(\int_0^\cdot Z_s \, dM_s\right) = p\,\mathcal{L}\left(\int_0^\cdot Z^1_s \, dM^1_s\right) + (1 - p)\,\mathcal{L}\left(\int_0^\cdot Z^2_s \, dM^2_s\right).
$$

The same technique can be applied to show that Equations (7), (8), (9), (10), (11), and (12) are convex constraints on $\mathcal{M}$. Thus $\mathcal{M}$ is convex.

$\square$

15

3 Construction of a weak solution

Theorem 3.1 *Assume that f satisfies hypotheses (H_1) and (H_2). Then Equation (1) admits a weak solution.*

This section is entirely devoted to the proof of Theorem 3.1, by constructing an extended solution to (1) in the terminology of Definition 2.4.

In Subsections 3.1 to 3.4, we only assume that f is measurable and satisfies the growth condition (H_1). Condition (H_2) will be needed only in Subsection 3.5, for the final part of the proof of Theorem 3.1.

Note that the counterexample given by Buckdahn and Engelbert in [9] does not fit in our framework, and we do not know any example of a BSDE of the form (2) or (1) under hypothesis (H_1) and (H_2) which has no strong solution.

3.1 Construction of an approximating sequence of solutions

Approximating equations The proof of Lemma 3.3 will show that (2) amounts to the following equations (16) and (17):

$$(16) \qquad Y_t = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_t^T f(s, X_s, Y_s, Z_s)\, ds\right)$$

$$\int_0^t Z_s\, dW_s = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_0^T f(s, X_s, Y_s, Z_s)\, ds\right)$$

$$(17) \qquad\qquad\qquad\qquad - \mathrm{E}\left(\xi + \int_0^T f(s, X_s, Y_s, Z_s)\, ds\right).$$

We can now write the approximating equations for (16) and (17):

$$(18) \qquad Y_t^{(n)} = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{t+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds\right)$$

$$\int_0^t Z_s^{(n)}\, dW_s = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds\right)$$

$$(19) \qquad\qquad\qquad\qquad - \mathrm{E}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds\right).$$

Here and in the sequel,

- f is extended by setting $f(t, x, y, z) = 0$ for $t > T$; similarly, for any function or process v defined on $[0, T]$, we set $v(t) = 0$ for $t > T$,

- we denote $\widetilde{Z}_s^{(n)} = \mathrm{E}^{\mathcal{F}_s}\left(Z_{s+1/n}^{(n)}\right)$.

Proposition 3.2 *The system* (18)-(19) *admits a unique strong solution* $(Y^{(n)}, Z^{(n)})$. *Furthermore, for every* $n \geq 1$, $Y_t^{(n)} \in \mathrm{L}_{\mathbb{R}^d}^2(\Omega)$ *for each* $t \in [0, T]$ *and* $Z^{(n)} \in \mathrm{L}_{\mathbb{L}}^2(\Omega \times [0, T])$.

Proof Let $T_k = T - \frac{k}{n}$, $k = 0, \ldots, \lceil nT \rceil$, where $\lceil nT \rceil$ is the integer part of nT. Observe first that for each k, (19) amounts on the interval $]T_{k+1}, T_k]$ to

(20) $$\int_{T_{k+1}}^t Z_s^{(n)}\, dW_s = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{T_{k+1}}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds\right)$$
$$- \mathrm{E}^{\mathcal{F}_{T_{k+1}}}\left(\xi + \int_{T_{k+1}}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds\right).$$

Now, the construction of $(Y^{(n)}, Z^{(n)})$ is easy by backward induction: For $T_1 \leq t \leq T = T_0$, we have $Y_t^{(n)} = \mathrm{E}^{\mathcal{F}_t}(\xi)$ and $(Z_t^{(n)})_{T_1 \leq t \leq T}$ is the unique predictable process such that $\mathrm{E}\int_{T_1}^T \left(Z_t^{(n)}\right)^2 ds < +\infty$ and

$$\int_{T_1}^t Z_s^{(n)}\, dW_s = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{T_1}^T f(s, X_s, Y_s^{(n)}, 0)\, ds\right)$$
$$- \mathrm{E}^{\mathcal{F}_{T_1}}\left(\xi + \int_{T_1}^T f(s, X_s, Y_s^{(n)}, 0)\, ds\right).$$

Suppose $(Y^{(n)}, Z^{(n)})$ is defined on the time interval $]T_k, T]$, with $k < \lceil nT \rceil$, then $Y^{(n)}$ is defined in a unique way on $]T_{k+1}, T_k]$ by (18) and then $Z^{(n)}$ on the same interval by (20). Furthermore, we get by induction from (20) that $Z^{(n)} \in \mathrm{L}_{\mathbb{L}}^2(\Omega \times [0, T])$. Then, using this latter result in (18), we deduce that $Y_t^{(n)} \in \mathrm{L}_{\mathbb{R}^d}^2(\Omega)$ for each $t \in [0, T]$. $\qquad\square$

The following result links (18) and (19) to an approximate version of (2):

Lemma 3.3 *Equations* (18) *and* (19) *are equivalent to*

(21) $$Y_t^{(n)} = \xi + \int_t^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds - \int_t^T Z_s^{(n)}\, dW_s - U_t^{(n)}$$

with $Y^{(n)}$ *adapted and*

$$U_t^{(n)} = \mathrm{E}^{\mathcal{F}_t}\left(\int_t^{t+1/n} f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds\right).$$

Proof Assume (18) and (19). Denoting

$$M_t^{(n)} = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds\right) = M_0^{(n)} + \int_0^t Z_s^{(n)}\,dW_s,$$

we get

$$M_t^{(n)} = \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{t+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds\right)$$
$$+ \int_0^t f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds + U_t^{(n)}.$$

By (18), this yields

$$M_t^{(n)} = Y_t^{(n)} + \int_0^t f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds + U_t^{(n)},$$

that is,

$$Y_t^{(n)} = M_t^{(n)} - \int_0^t f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds - U_t^{(n)}$$
$$= M_0^{(n)} + \int_0^t Z_s^{(n)}\,dW_s - \int_0^t f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds - U_t^{(n)}.$$

In particular,

$$Y_T^{(n)} = \xi = M_0^{(n)} + \int_0^T Z_s^{(n)}\,dW_s - \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds$$

thus

$$Y_t^{(n)} - Y_T^{(n)} = -\int_t^T Z_s^{(n)}\,dW_s + \int_t^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds - U_t^{(n)},$$

which proves (21).

Conversely, assume (21) and that $Y^{(n)}$ is adapted. Denote $V_t^{(n)} = \int_0^t Z_s^{(n)} \, dW_s$. We have

$$Y_t^{(n)} = \mathrm{E}^{\mathcal{F}_t}\left(Y_t^{(n)}\right)$$

$$= \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_t^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \int_t^T Z_s^{(n)} \, dW_s\right.$$

$$\left. - \int_t^{t+1/n} f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds\right)$$

$$= \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{t+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds\right) - \mathrm{E}^{\mathcal{F}_t}\left(V_T^{(n)} - V_t^{(n)}\right)$$

$$= \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{t+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds\right),$$

which proves (18).

Now, using (18) and (21), we have

$$\mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds\right)$$

$$= \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{t+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds\right)$$

$$+ \int_0^t f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds + U_t^{(n)}$$

$$= Y_t^{(n)} + \int_0^t f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds + U_t^{(n)}$$

$$= \xi + \int_t^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \int_t^T Z_s^{(n)} \, dW_s - U_t^{(n)}$$

$$+ \int_0^t f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds + U_t^{(n)}$$

$$= \xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \int_t^T Z_s^{(n)} \, dW_s.$$

In particular,

$$\mathrm{E}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds\right)$$

$$= \xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \int_0^T Z_s^{(n)} \, dW_s.$$

19

Thus

$$\int_0^t Z_s^{(n)} \, dW_s = \int_0^T Z_s^{(n)} \, dW_s - \int_t^T Z_s^{(n)} \, dW_s$$

$$= \left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \mathrm{E}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right) \right)$$

$$- \left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right) \right)$$

$$= \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right) - \mathrm{E}\left(\xi + \int_0^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right)$$

which proves (19). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

3.2 Boundedness and continuity results

In this part, we show some results that will be useful to prove the relative compactness in distribution of the sequence $\left(Y^{(n)}, \int_{\cdot}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds, \int_{\cdot}^T Z_s^{(n)} \, dW_s, Z^{(n)} \right)$ in some properly chosen state space.

Lemma 3.4 *Let*

$$\widetilde{Y}_t^{(n)} = Y_t^{(n)} + U_t^{(n)} = \xi + \int_t^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \int_t^T Z_s^{(n)} \, dW_s.$$

There exist constants $\mathfrak{a}, \mathfrak{b} > 0$ such that, for all t such that $0 \le t \le T$,

$$(22) \qquad\qquad \mathrm{E} \int_t^T \left\| Z_s^{(n)} \right\|^2 ds \le \mathfrak{a} \, \mathrm{E} \int_t^T \left\| \widetilde{Y}_s^{(n)} \right\|^2 ds + \mathfrak{b}.$$

Proof Applying Itô's formula to the semi-martingale $\left\| \widetilde{Y}_t^{(n)} \right\|^2$, taking expectation of both sides and using the fact that $t' \mapsto \int_t^{t'} \left\langle \widetilde{Y}_t^{(n)}, Z_s^{(n)} \, dW_s \right\rangle$ is a martingale (thanks to Proposition 3.2), we get

$$\mathrm{E} \left\| \widetilde{Y}_t^{(n)} \right\|^2 = \mathrm{E} \left\| \xi \right\|^2 + 2 \, \mathrm{E} \int_t^T \widetilde{Y}_s^{(n)} . f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) ds - \mathrm{E} \int_t^T \left\| Z_s^{(n)} \right\|^2 ds.$$

Thus

$$\mathrm{E} \int_t^T \left\| Z_s^{(n)} \right\|^2 ds \le \mathrm{E} \left\| \xi \right\|^2 + 2 \, \mathrm{E} \int_t^T \left\| \widetilde{Y}_s^{(n)} \right\| . \left\| f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \right\| ds.$$

20

From (H_1), this entails

$$\mathrm{E}\int_t^T \left\|Z_s^{(n)}\right\|^2 ds \le \mathrm{E}\left\|\xi\right\|^2 + 2C_f\,\mathrm{E}\int_t^T \left\|\widetilde{Y}_s^{(n)}\right\|(1+\left\|\widetilde{Z}_s^{(n)}\right\|)\,ds.$$

Using that, for $a \ge 0, b \ge 0$, and $\lambda \ne 0$, we have $2ab \le a^2\lambda^2 + b^2/\lambda^2$, we get

$$\mathrm{E}\int_t^T \left\|\widetilde{Y}_s^{(n)}\right\|(1+\left\|\widetilde{Z}_s^{(n)}\right\|)\,ds$$

$$\le 2\lambda^2\,\mathrm{E}\int_t^T \left\|\widetilde{Y}_s^{(n)}\right\|^2 ds + (T-t)/\lambda^2 + 1/\lambda^2\,\mathrm{E}\int_t^T \left\|\widetilde{Z}_s^{(n)}\right\|^2 ds$$

$$\le 2\lambda^2\,\mathrm{E}\int_t^T \left\|\widetilde{Y}_s^{(n)}\right\|^2 ds + (T-t)/\lambda^2 + 1/\lambda^2\,\mathrm{E}\int_t^T \left\|Z_s^{(n)}\right\|^2 ds.$$

Thus, taking $\lambda^2 > 2C_f$,

$$(1-2C_f/\lambda^2)\,\mathrm{E}\int_t^T \left\|Z_t^{(n)}\right\|^2 ds \le \mathrm{E}\left\|\xi\right\|^2 + 2C_f\left(T/\lambda^2 + 2\lambda^2\,\mathrm{E}\int_t^T \left\|\widetilde{Y}_t^{(n)}\right\|^2 ds\right)$$

which yields (22). $\qquad\square$

Proposition 3.5 *Let* $\widetilde{Y}_t^{(n)} = Y_t^{(n)} + U_t^{(n)}$ *be as in Lemma 3.4. The families* $(\widetilde{Y}_t^{(n)})_{0\le t\le T,\,n\ge 1}$, $(Y_t^{(n)})_{0\le t\le T,\,n\ge 1}$ *and* $(U_t^{(n)})_{0\le t\le T,\,n\ge 1}$ *are bounded in* $\mathrm{L}^2_{\mathbb{R}^d}(\Omega)$.

Proof We have

$$\widetilde{Y}_t^{(n)} = Y_t^{(n)} + U_t^{(n)}$$

$$= \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{t+1/n}^T f(s,X_s,Y_s^{(n)},\widetilde{Z}_s^{(n)})\right) + \mathrm{E}^{\mathcal{F}_t}\left(\int_t^{t+1/n} f(s,X_s,Y_s^{(n)},\widetilde{Z}_s^{(n)})\,ds\right)$$

$$= \mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_t^T f(s,X_s,Y_s^{(n)},\widetilde{Z}_s^{(n)})\,ds\right).$$

We deduce the following inequalities, where C denotes some constant which is not necessarily the same at each line but does not depend on n:

$$\mathrm{E}\left\|\widetilde{Y}_t^{(n)}\right\|^2 = \mathrm{E}\left\|\mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_t^T f(s,X_s,Y_s^{(n)},\widetilde{Z}_s^{(n)})\,ds\right)\right\|^2$$

$$\le C\,\mathrm{E}\left(\left\|\xi\right\|^2 + \int_t^T (1+\left\|Z_s^{(n)}\right\|^2)\,ds\right)$$

$$\le C\left(1 + \int_t^T \mathrm{E}\left\|\widetilde{Y}_s^{(n)}\right\|^2 ds\right).$$

21

The last inequality is a consequence of Lemma 3.4. Let $g(t) = \mathrm{E}\left\|\widetilde{Y}_{T-t}^{(n)}\right\|^2$. The preceding inequalities yield

$$g(t) \leq C\left(1 + \int_0^t g(s)\,ds\right).$$

Thus, by Gronwall's Lemma,

$$g(t) \leq C\left(1 + C\int_0^t e^{C(t-s)}\,ds\right) \leq C\left(1 + C\int_0^T e^{C(T-s)}\,ds\right)$$

which proves that $(\widetilde{Y}_t^{(n)})_{0 \leq t \leq T, n \geq 1}$ is bounded in $\mathrm{L}^2_{\mathbb{R}^d}(\Omega)$.

Now, we have, using again Lemma 3.4,

$$\mathrm{E}\left(\left\|Y_t^{(n)}\right\|^2\right) = \mathrm{E}\left\|\mathrm{E}^{\mathcal{F}_t}\left(\xi + \int_{t+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\,ds\right)\right\|^2$$

$$\leq C\,\mathrm{E}\left(\|\xi\|^2 + \int_{t+1/n}^T (1 + \left\|Z_s^{(n)}\right\|^2)\,ds\right)$$

$$\leq C\left(1 + \int_0^t \mathrm{E}\left\|\widetilde{Y}_s^{(n)}\right\|^2\,ds\right)$$

which proves that $(Y_t^{(n)})_{0 \leq t \leq T, n \geq 1}$ is bounded in $\mathrm{L}^2_{\mathbb{R}^d}(\Omega)$.

The boundedness in $\mathrm{L}^2_{\mathbb{R}^d}(\Omega)$ of $(U_t^{(n)})_{0 \leq t \leq T, n \geq 1}$ follows immediately from that of $(\widetilde{Y}_t^{(n)})_{0 \leq t \leq T, n \geq 1}$ and $(Y_t^{(n)})_{0 \leq t \leq T, n \geq 1}$. $\qquad\square$

Corollary 3.6 *The sequences $(Z^{(n)})_{n \geq 1}$ and $(\widetilde{Z}^{(n)})_{n \geq 1}$ are bounded in $\mathrm{L}^2_{\mathbb{L}}(\Omega \times [0, T])$, and we have*

$$(23) \qquad \sup_{n \geq 1} \mathrm{E}\left(\sup_{0 \leq t \leq T}\left\|\int_t^T Z_s^{(n)}\,dW_s\right\|^2\right) < +\infty.$$

Proof The boundedness of $(Z^{(n)})_{n \geq 1}$ and $(\widetilde{Z}^{(n)})_{n \geq 1}$ is a direct consequence of Lemma 3.4 and Proposition 3.5. Then (23) follows by Itô's isometry, Doob's inequality, and the fact that

$$\left\|\int_t^T Z_s^{(n)}\,dW_s\right\| \leq \left\|\int_0^T Z_s^{(n)}\,dW_s\right\| + \left\|\int_0^t Z_s^{(n)}\,dW_s\right\|.$$

$\qquad\square$

Lemma 3.7 *Let $1 \le q < 2$. We have*

(24)
$$\lim_{n \to \infty} \mathrm{E}\left(\sup_{0 \le t \le T}\left\|U_t^{(n)}\right\|^q\right) = 0.$$

Proof For each n, we can find an $\mathcal{F}_T$-measurable time τ_n such that

$$\sup_{0 \le t \le T}\int_t^{t+1/n}\left(1+\left\|Z_s^{(n)}\right\|\right)^q ds = \int_{\tau_n}^{\tau_n+1/n}\left(1+\left\|Z_s^{(n)}\right\|\right)^q ds.$$

Let $\mathfrak{M}_2 = \sup_n \mathrm{E}\int_0^T\left(1+\left\|Z_s^{(n)}\right\|\right)^2 ds$. By Corollary 3.6, we have $\mathfrak{M}_2 < +\infty$. Let q' such that $q < q' < 2$. Using the growth condition (H_1) and Doob's inequality applied to the martingale $\mathrm{E}^{\mathcal{F}_t}\left(\int_{\tau_n}^{\tau_n+1/n}\left(1+\left\|Z_s^{(n)}\right\|\right)^q ds\right)$, we get

$$\mathrm{E}\left(\sup_{0 \le t \le T}\left\|U_t^{(n)}\right\|^q\right)$$

$$\le C_f^q\,\mathrm{E}\left(\sup_{0 \le t \le T}\mathrm{E}^{\mathcal{F}_t}\left(\int_t^{t+1/n}\left(1+\left\|Z_s^{(n)}\right\|\right)^q ds\right)\right)$$

$$\le C_f^q\,\mathrm{E}\left(\sup_{0 \le t \le T}\mathrm{E}^{\mathcal{F}_t}\left(\int_{\tau_n}^{\tau_n+1/n}\left(1+\left\|Z_s^{(n)}\right\|\right)^q ds\right)\right)$$

$$\le C_f^q\left(\mathrm{E}\left(\sup_{0 \le t \le T}\mathrm{E}^{\mathcal{F}_t}\left(\int_{\tau_n}^{\tau_n+1/n}\left(1+\left\|Z_s^{(n)}\right\|\right)^q ds\right)^{q'/q}\right)\right)^{q/q'}$$

$$\le \frac{q'}{q'-q}\,C_f^q\left(\mathrm{E}\left(\int_{\tau_n}^{\tau_n+1/n}\left(1+\left\|Z_s^{(n)}\right\|\right)^{q'} ds\right)\right)^{q/q'}$$

$$\le \frac{q'}{q'-q}\,C_f^q\left(\frac{1}{n}\right)^{(2-q')/2}\left(\left(\mathrm{E}\int_0^T\left(1+\left\|Z_s^{(n)}\right\|\right)^2 ds\right)^{q'/2}\right)^{q/q'}$$

$$= \frac{q'}{q'-q}\,C_f^q\left(\frac{1}{n}\right)^{(2-q')/2}\left(\mathrm{E}\int_0^T\left(1+\left\|Z_s^{(n)}\right\|\right)^2 ds\right)^{q/2}$$

$$\le \frac{q'}{q'-q}\,C_f^q\,\mathfrak{M}_2^{q/2}\left(\frac{1}{n}\right)^{(2-q')/2},$$

which proves (24). $\qquad\square$

Lemma 3.8 *We have*

$$\sup_{n \geq 1} \mathrm{E} \left(\sup_{0 \leq t \leq T} \left\| Y_t^{(n)} \right\|^2 \right) < +\infty.$$

Proof Using (21), we get

$$\sup_{0 \leq t \leq T} \left\| Y_t^{(n)} \right\|^2 \leq A_n + B_n + C_n$$

where

$$A_n = 3 \sup_{0 \leq t \leq T} \left\| \xi + \int_{t+1/n}^{T} f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right\|^2,$$

$$B_n = 3 \sup_{0 \leq t \leq T} \left\| \int_t^T Z_s^{(n)} \, dW_s \right\|^2,$$

$$C_n = 3 \sup_{0 \leq t \leq T} \left\| U_t^{(n)} \right\|^2.$$

By Corollary 3.6, $(Z^{(n)})_{n \geq 1}$ is bounded in $\mathrm{L}_{\mathbb{L}}^2(\Omega \times [0,T])$, thus using the growth condition (H_1), we get

$$\sup_n \mathrm{E} \left(\sup_{0 \leq t \leq T} \left(\|\xi\|^2 + C_f^2 \int_{t+1/n}^T \left(1 + \left\| Z_s^{(n)} \right\| \right)^2 ds \right) \right) < +\infty$$

which entails $\sup_n \mathrm{E}(A_n) < +\infty$. On the other hand, $V_t^{(n)} := \int_0^t Z_s^{(n)} \, dW_s$ is a martingale, so, using again Corollary 3.6,

$$\sup_n \mathrm{E}(B_n) \leq C \sup_n \mathrm{E} \left\| V_T^{(n)} \right\|^2 < +\infty.$$

Finally from (H_1) and the boundedness of $(Z^{(n)})_{n \geq 1}$ in $\mathrm{L}_{\mathbb{L}}^2(\Omega \times [0,T])$ (see Corollary 3.6), we have

$$\sup_n \mathrm{E}(C_n) \leq 3 \sup_{0 \leq t \leq T} C_f^2 \, \mathrm{E} \left(\int_0^T \left(1 + \left\| Z_s^{(n)} \right\| \right)^2 ds \right) < +\infty.$$

$\square$

3.3 Compactness results

Lemma 3.9 *The sequence* $(\int_{\cdot}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds)_{n \geq 1}$ *is tight in* $C_{\mathbb{R}^d}[0, T]$.

Proof Let us denote $\Sigma^{(n)} = \int_{\cdot}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds$. By a criterion of Aldous [1, 18], we only need to prove that

$$(A) \qquad \forall \epsilon > 0,\ \exists R > 0, \forall n \geq 1,\ P\left(\sup_{0 \leq t \leq T} \left\| \Sigma_t^{(n)} \right\| \geq R \right) \leq \epsilon$$

$$(B) \quad \forall \epsilon > 0, \forall \eta > 0,\ \exists \delta > 0 : \forall n \geq 1, \quad \sup_{\substack{\sigma, \tau \in \mathfrak{T} \\ 0 \leq |\tau - \sigma| \leq \delta}} P\left(\left\| \Sigma_\tau^{(n)} - \Sigma_\sigma^{(n)} \right\| \geq \eta \right) \leq \epsilon$$

where $\mathfrak{T}$ denotes the set of stopping times with values in $[0, T]$. We are going to prove the slightly stronger properties

$$(25) \qquad\qquad \sup_{n \geq 1} E\left(\sup_{0 \leq t \leq T} \left\| \Sigma_t^{(n)} \right\| \right) < +\infty,$$

$$(26) \qquad \forall \epsilon > 0,\ \exists \delta > 0 : \sup_{n \geq 1} \sup_{\substack{\sigma, \tau \in \mathfrak{T} \\ |\tau - \sigma| \leq \delta}} E\left\| \Sigma_\tau^{(n)} - \Sigma_\sigma^{(n)} \right\| < +\epsilon.$$

As

$$\Sigma_t^{(n)} = Y_t^{(n)} - \xi + \int_t^T Z_s^{(n)}\, dW_s + U_t^{(n)},$$

we can, for example, deduce (25) from Corollary 3.6, Lemma 3.7, and Lemma 3.8.

Now, let $\sigma, \tau \in \mathfrak{T}$, with $|\tau - \sigma| \leq \delta$. We have $E\left\| \Sigma_\tau^{(n)} - \Sigma_\sigma^{(n)} \right\| = E\left\| \Sigma_{\sigma \vee \tau}^{(n)} - \Sigma_{\sigma \wedge \tau}^{(n)} \right\|$. Thus we can assume without loss of generality that $\sigma \leq \tau$. Then

$$E\left\| \Sigma_\tau^{(n)} - \Sigma_\sigma^{(n)} \right\| = E\left\| \int_\sigma^\tau f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds \right\|,$$

$$\leq (E(\tau - \sigma))^{1/2} \left(E \int_0^T \left\| f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \right\|^2 ds \right)^{1/2}$$

$$\leq \delta^{1/2} C_f \left(E \int_0^T (1 + \left\| Z_s^{(n)} \right\|)^2\, ds \right)^{1/2}$$

and (26) follows from Corollary 3.6. $\qquad \square$

The topology S and Condition UT   In order to prove the tightness of $(Y^{(n)})_{n \geq 1}$, we will use Meyer-Zheng criterion [29] and Jakubowski's topology S [21] on the space $\mathbb{D} := \mathbb{D}_{\mathbb{R}^d}[0, T]$. First, we need some definitions.

Let $\mathbb{V} \subset \mathbb{D}$ be the subspace of elements of $\mathbb{D}$ which have finite variation. The topology S on $\mathbb{D}$ is defined by its convergent sequences: A sequence (x_n) in $\mathbb{D}$ converges for S to a limit $x \in \mathbb{D}$ if, from any subsequence of (x_n), one can extract a further subsequence (x'_n) such that, for every $\epsilon > 0$, there exist a sequence $(v_{n,\epsilon})$ of elements of $\mathbb{V}$ and $v_\epsilon \in \mathbb{V}$ (depending on the subsequence (x'_n)) such that

(i) $\sup_n \sup_{t \in [0,T]} \|x'_n(t) - v_{n,\epsilon}(t)\| \leq \epsilon$ and $\sup_{t \in [0,T]} \|x(t) - v_\epsilon(t)\| \leq \epsilon$,

(ii) $\lim_{n \to \infty} \int_0^T f(t)\, dv_{n,\epsilon}(t) = \int_0^T f(t)\, dv_\epsilon(t)$ for every continuous function f defined on $[0, T]$.

We denote $\mathbb{D}_S$ the space $\mathbb{D}$ endowed with S. The topology S is coarser than Skorokhod's topology J_1, which is Polish, thus S is Lusin (see [33] on properties of Lusin spaces). In particular, by [33, Corollary 2 page 101], S has the same Borel sets as J_1, thus the Borel subsets of S are generated by the projection mappings $\pi_t : x \mapsto x(t)$ for $t \in [0, T]$). Furthermore, S is finer than the Meyer-Zheng topology [29], which is the topology on $\mathbb{V}_{\mathbb{R}^d}$ induced by $L^0_{\mathbb{R}^d}([0, T], dt)$. In particular, S is (separably) submetrizable, that is, there exists a (separable) metrizable topology which is coarser than S. Equivalently, one can find a countable set of S-continuous real-valued functions which separate the points of $\mathbb{D}$. This implies that S is Hausdorff and that the compact subsets of $\mathbb{D}_S$ are metrizable.

Another important feature of S is that the addition $(x, y) \mapsto x + y$ is S-sequentially continuous on $\mathbb{D}_S \times \mathbb{D}_S$.

A criterion of tightness on $\mathbb{D}_S$ is the so-called condition UT (see [21, Theorem 4.2]): Let $\mathcal{H}$ denote the set of elementary real valued predictable processes bounded by 1, i.e. processes of the form

$$H_t = \mathbb{1}_{[t_0, t_1]}(t) H_0 + \mathbb{1}_{]t_1, t_2]}(t) H_1 + \cdots + \mathbb{1}_{]t_{n-1}, t_n]}(t) H_{t_{n-1}}$$

where $0 = t_0 \leq \cdots \leq t_n \leq T$ and each H_i is bounded by 1 and $\mathcal{F}_{t_i}$-measurable. Let $(K^\alpha)_{\alpha \in A}$ be a family of $\mathbb{D}$-valued processes. We say that (K^α) *satisfies Condition UT* if the family of all stochastic integrals $\int H\, dK^\alpha$, where $\alpha \in A$ and $H \in \mathcal{H}$, is uniformly tight. Condition UT was considered for the first time by Stricker [35], to prove compactness in the Meyer-Zheng topology. Discussions on this condition can be found in [22, 28].

We now consider a stronger condition, proposed by Meyer and Zheng [29]: Let K be an adapted process defined on the time interval $[0, T]$, with values in $\mathbb{R}^d$. For any finite partition $\pi = (t_0, \dots, t_n)$ of $[0, T]$, let us denote

$$\mathfrak{N}_\pi(K) = \mathrm{E} \, \|K_T\| + \sum_{i=0}^{n-1} \left\| \mathrm{E}^{\mathcal{F}_{t_i}} \left(K_{t_{i+1}} - K_{t_i} \right) \right\|,$$

and define the *conditional variation* $\mathfrak{N}(K)$ of K by

$$\mathfrak{N}(K) = \sup_\pi \mathfrak{N}_\pi(K).$$

By [35, Théorème 3], if a family (K^α) of adapted $\mathbb{D}$-valued processes satisfies

$$\sup_\alpha \mathfrak{N}(K^\alpha) < \infty,$$

then Condition UT holds for (K^α).

An adapted stochastic process K such that $\mathfrak{N}(K) < \infty$ is called a *quasimartingale*. Let us mention that, if the quasimartingale K is right-continuous in probability, then it has a càdlàg adapted version (assuming the right-continuity of $(\mathcal{F}_t)$), see [8, Theorem 4.1].

Proposition 3.10 *The sequences $(Y^{(n)})_{n \geq 1}$ and $(\int_{\cdot}^T Z_s^{(n)} \, dW_s)_{n \geq 1}$ are tight sequences of $\mathbb{D}$-valued random variables, for the topology S.*

Proof First, we need to check that, for each integer $n \geq 1$, the process

$$(27) \qquad Y_t^{(n)} = \mathrm{E}^{\mathcal{F}_t} \left(\xi + \int_{t+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right)$$

has a $\mathbb{D}$-valued version. Let us prove that it is continuous in L^1 and a quasimartingale. As $(\mathcal{F}_t)$ is a Brownian filtration, the martingale

$$t \mapsto \mathrm{E}^{\mathcal{F}_t} \left(\xi + \int_{r+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right)$$

has a continuous version for each fixed $r \in [0, T - 1/n]$, thus it is continuous in L^1, i.e. the mapping
(28)

$$\begin{cases} [0,T] \times [0, T - 1/n] & \to \quad \mathrm{L}^1 \\ (t, r) & \mapsto \quad \mathrm{E}^{\mathcal{F}_t} \left(\xi + \int_{r+1/n}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right) \end{cases}$$

is continuous in the variable t. On the other hand, we have, for fixed $t \in [0, T]$ and for $0 \le r_1 \le r_2 \le T - 1/n$ such that $r_2 - r_1 \le 1/n$,

$$
\left\| \mathrm{E}^{\mathcal{F}_t} \left(\int_{r_1+1/n}^{r_2+1/n} f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right) \right\|
$$

$$
\le C_f \, \mathrm{E}^{\mathcal{F}_t} \left(\int_{r_1+1/n}^{r_2+1/n} (1 + \left\| \widetilde{Z}_s^{(n)} \right\|) \, ds \right)
$$

$$
\le (r_2 - r_1)^{1/2} C_f \left(\mathrm{E}^{\mathcal{F}_t} \left(\int_0^T (1 + \left\| \widetilde{Z}_s^{(n)} \right\|)^2 \, ds \right) \right)^{1/2}.
$$

Therefore, by Corollary 3.6, the mapping (28) is continuous in r uniformly with respect to t, thus it is jointly continuous, which proves the continuity in L^1 of the process (27) for each $n \ge 1$.

Now, we have, for any subdivision $\pi = (t_0, \ldots, t_m)$ of $[0, T]$,

$$
\sup_n \mathfrak{N}_\pi(Y^{(n)}) = \sup_n \mathrm{E} \left(\|\xi\| + \sum_{i=0}^{m-1} \left\| \mathrm{E}^{\mathcal{F}_{t_i}} \left(Y_{t_{i+1}}^{(n)} - Y_{t_i}^{(n)} \right) \right\| \right)
$$

$$
\le \sup_n \mathrm{E} \left(\|\xi\| + \sum_{i=0}^{m-1} \left\| \int_{t_i+1/n}^{t_{i+1}+1/n} f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right\| \right)
$$

$$
\le \sup_n \mathrm{E} \left(\|\xi\| + \int_0^T \left\| f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \right\| \, ds \right).
$$

This estimation does not depend on π, thus, using Corollary 3.6,

$$
\sup_n \mathfrak{N}(Y^{(n)}) \le \mathrm{E} \left(\|\xi\| \right) + \sup_n \mathrm{E} \left(\int_0^T C_f(1 + \left\| \widetilde{Z}_s^{(n)} \right\|) \, ds \right)
$$

$$
\le \mathrm{E} \left(\|\xi\| \right) + \sup_n C_f \left(T + T^{1/2} \left(\mathrm{E} \left(\int_0^T \left\| \widetilde{Z}_s^{(n)} \right\|^2 \right) \, ds \right)^{1/2} \right)
$$

$$
< +\infty.
$$

This proves that each $Y^{(n)}$ is a quasimartingale, and that the sequence $(Y^{(n)})_{n \ge 1}$ satisfies Condition UT. Furthermore, for each $n \ge 1$, as $Y^{(n)}$ is right-continuous in L^1, it has a càdlàg version, thanks to [8, Theorem 4.1]. Thus, by [21, Theorem 4.2], the sequence $(Y^{(n)})_{n \ge 1}$ is tight in $\mathbb{D}_\mathrm{S}$.

Similarly,

$$\sup_n \mathfrak{N}_\pi \left(\int_0^\cdot Z_s^{(n)} \, dW_s \right)$$

$$= \sup_n \mathrm{E}\left(\left\| \int_0^T Z_s^{(n)} \, dW_s \right\| + \sum_{i=0}^{m-1} \left\| \mathrm{E}^{\mathcal{F}_{t_i}} \left(\int_{t_i}^{t_{i+1}} Z_s^{(n)} \, dW_s \right) \right\| \right)$$

$$\leq 2 \sup_n \mathrm{E} \left\| \int_0^T Z_s^{(n)} \, dW_s \right\|,$$

thus

$$\sup_n \mathfrak{N}\left(\int_0^\cdot Z_s^{(n)} \, dW_s \right) \leq 2 \sup_n \mathrm{E} \left\| \int_0^T Z_s^{(n)} \, dW_s \right\| < +\infty.$$

Thus $(\int_0^\cdot Z_s^{(n)} \, dW_s)_{n \geq 1}$ satisfies Condition UT. (Alternatively, if $(M^{(n)})$ is a sequence of $\mathbb{R}^d$-valued martingales on the time interval $[0, T]$ such that $(M_T^{(n)})$ is uniformly bounded in L^2, it is well known and easy to check that $(M^{(n)})$ satisfies Condition UT.) Again by [21, Theorem 4.2], this proves that $(\int_0^\cdot Z_s^{(n)} \, dW_s)_{n \geq 1}$ is tight in $\mathbb{D}_\mathrm{S}$. Finaly, it is straightforward to check that the mapping

$$\begin{cases} \mathbb{D} & \to & \mathbb{D} \\ u & \mapsto & u(T) - u \end{cases}$$

is sequentially continuous for the topology S, thus $(\int_\cdot^T Z_s^{(n)} \, dW_s)_{n \geq 1}$ is tight in $\mathbb{D}_\mathrm{S}$. $\qquad\square$

3.4 Construction of a weak limit process

This part of the construction of a weak solution follows the same lines as in [23], with some complications due to the processes $Z^{(n)}$.

Young measures Let us recall the definition and main properties of Young measures, see [37, 7] for introductions to the topic, and [12] for the setting of nonnecessarily regular topological spaces, which we need here. Let $\mathbb{E}$ be a Suslin topological space (i.e. $\mathbb{E}$ is a Hausdorff topological space and there exists a Polish space $\mathbb{S}$ and a continuous surjective mapping from $\mathbb{S}$ onto $\mathbb{E}$, see [33] for the properties of Suslin spaces, or [12, Chapter 1] for a survey without proofs). Let $\mathcal{B}(\mathbb{E})$ be the Borel σ-algebra of $\mathbb{E}$. A *Young measure μ with basis* P *on* $\mathbb{E}$ is a probability measure on $\Omega \times \mathbb{E}$, such that

for any set $A \in \mathcal{F}$, $\mu(A \times \mathbb{E}) = P(A)$. The space of Young measures with basis P is denoted by $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$. It is very useful to describe a Young measure μ by its *disintegration* (μ_ω) with respect to P (see Definition 2.8). The space $L^0(\Omega, \mathcal{F}, P; \mathbb{E})$ of measurable functions from Ω to $\mathbb{E}$ is embedded in $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$ in the following way: we identify every $u \in L^0(\Omega, \mathcal{F}, P; \mathbb{E})$ with the Young measure $\delta_{u(\omega)} \otimes d P(\omega)$, where $\delta_{u(\omega)}$ denotes the Dirac mass at $u(\omega)$. In other words, u is identified with the unique Young measure μ whose support is the graph of u. The set $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$ is endowed with a topology defined as follows: A generalized sequence[2] (μ^α) of Young measures converges to a Young measure μ if, for each bounded measurable $\Phi : \Omega \times \mathbb{E} \to \mathbb{R}$ such that $\Phi(\omega, .)$ is continuous for all $\omega \in \Omega$, the generalized sequence $(\mu^\alpha(\Phi))$ converges to $\mu(\Phi)$. In this case, we say that (μ^α) *converges stably*, or $\mathcal{F}$-*stably*, to μ (this terminology stems from Rényi [32]).

Note that the restriction of the topology of stable convergence to $L^0(\Omega; \mathbb{E})$ is the topology of convergence in probability, see [37, 12].

We say that a subset $\mathcal{K}$ of $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$ is *tight* if, for each $\epsilon > 0$, there exists a compact subset K of $\mathbb{E}$ such that $\inf_{\mu \in \mathcal{K}} \mu(\Omega \times K) \geq 1 - \epsilon$. In the case when $\mathcal{K} \subset L^0(\Omega, \mathcal{F}, P; \mathbb{E})$, this is the usual tightness notion for random variables. By [12, Theorem 4.3.5], if the compact subsets of $\mathbb{E}$ are metrizable, and if $\mathcal{K}$ is tight, then $\mathcal{K}$ is relatively compact and relatively sequentially compact in $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$. The converse is true if $\mathbb{E}$ has the Prohorov property.

We will need a result on convergence of Young measures with respect to sequentially continuous integrands:

Lemma 3.11 *Assume that $\mathbb{E}$ is a Suslin submetrizable topological space. Let (μ^n) be a tight sequence in $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$ which stably converges to a Young measure μ. Let $f : \Omega \times \mathbb{E} \to \mathbb{R}$ be a bounded measurable function such that $f(\omega, .)$ is* sequentially *continuous for each $\omega \in \Omega$. Then $\lim_n \mu^n(f) = \mu(f)$.*

Proof By Balder's extension of Komlós Theorem for Young measures [5, 6] which is valid for Hausdorff spaces with metrizable compact subsets [12, Lemma 4.5.4], we can extract from every subsequence of (μ^n) a further subsequence (which we still denote by (μ^n) for simplicity of notations), which

---

[2]see [24] on generalized sequences, also called nets, however we do not need them in the sequel, because we use sequential compactness results. Note also that, when $\mathbb{E}$ is metrizable, the space $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$ is metrizable too, and we can characterize its topology using convergent sequences instead of convergent generalized sequences.

K–converges to μ, that is, for each subsequence (ν^n) of (μ^n), we have

$$(29) \qquad \lim_n \frac{1}{n} \sum_{k=1}^{n} \nu_\omega^n = \mu_\omega \text{ a.e.}$$

where the limit is taken in the narrow convergence, i.e. $\lim_n \frac{1}{n} \sum_{k=1}^{n} \nu_\omega^n(g) = \mu_\omega(g)$ for every bounded continuous $g : \mathbb{E} \to \mathbb{R}$. Let us denote $\lambda^n = \frac{1}{n} \sum_{k=1}^{n} \nu^n$, and let us prove that

$$(30) \qquad \lim_n \int_{\mathbb{E}} f(\omega, x) \, d\lambda_\omega^n(x) = \int_{\mathbb{E}} f(\omega, x) \, d\mu_\omega(x) \text{ a.e.}$$

Let ω be in the almost sure set on which the convergence in (29) holds. As $\mathbb{E}$ admits a coarser separable metrizable topology, we can apply Jakubowski's extension of Skorokhod's representation theorem [20]: for every subsequence of (λ_ω^n), we can find a further subsequence $(\lambda_\omega^{n_k})$ (which depends on ω), a probability space $(\Omega', \mathcal{F}', \mathrm{P}')$, and random $\mathbb{E}$-valued variables $X_1, \ldots, X_k, \ldots$ and X defined on Ω' such that the law of X_k is $\lambda_\omega^{n_k}$ for each k, the law of X is μ_ω, and (X_k) converges P'-a.e. to X. For such an ω, we have, by the dominated convergence theorem,

$$\lim_k \int_{\mathbb{E}} f(\omega, x) \, d\lambda_\omega^{n_k}(x) = \lim_k \int_{\Omega'} f(\omega, X_k) \, d\mathrm{P}'$$
$$= \int_{\Omega'} f(\omega, X) \, d\mathrm{P}' = \int_{\mathbb{E}} f(\omega, x) \, d\mu_\omega(x).$$

Thus, for ω in the almost sure set of (29), every subsequence of $(\lambda_\omega^{n_k})$ has a further subsequence for which the convergence in (30) holds. This proves (30). We deduce that, for any subsequence of (μ^n) we can extract a further subsequence (ν^n) such that

$$(31) \qquad \lim_n \frac{1}{n} \sum_{k=1}^{n} \nu^n(f) = \mu(f),$$

which proves the lemma. $\qquad\qquad\qquad\qquad\qquad$ $\square$

The following technical lemma will be useful for limits of integrals of unbounded integrands with respect to Young measures.

Lemma 3.12 *Let $\mathbb{E}$ be a Suslin submetrizable topological space, and let (X_n) be a sequence of $\mathbb{E}$-valued random variables defined on Ω. Assume that (X_n) stably converges to a Young measure $\mu \in \mathcal{Y}(\Omega, \mathcal{F}, \mathrm{P}; \mathbb{E})$ (where each X_n is identified with the Young measure $\delta_{X_n(\omega)} \otimes d\mathrm{P}(\omega)$). Let $\Phi : \Omega \times \mathbb{E} \to \mathbb{R}$ be measurable such that*

31

(i) $\Phi(\omega, .)$ is sequentially continuous for all $\omega \in \Omega$,

(ii) The sequence $(\Phi(., X_n))$ is uniformly integrable.

Then Φ is μ-integrable, and

$$\lim_n \mathrm{E}\, \Phi(., X_n) = \int_{\Omega \times \mathbb{E}} \Phi \, d\mu.$$

Proof We only need to prove Lemma 3.12 in the case when $\Phi \geq 0$, the general result comes from $\Phi = \Phi_+ - \Phi_-$.

For each $N \geq 0$, we have

(32)
$$\lim_{N \to +\infty} \sup_n \mathrm{E}\left(\Phi(., X_n)\, \mathbb{1}_{\{\Phi(., X_n) \geq N\}}\right) = 0.$$

Set

$$\Phi^N = \begin{cases} \Phi & \text{if } \Phi \leq N \\ N & \text{if } \Phi \geq N. \end{cases}$$

From the definition of stable convergence and Lemma 3.11, we have, for each N,

(33)
$$\lim_{n \to +\infty} \mathrm{E}\, \Phi^N(., X_n) = \mu(\Phi^N).$$

Furthermore, by (32), the convergence in (33) is uniform with respect to N. We thus have, with the help of Beppo Levi's lemma:

$$\begin{aligned}
\mu(\Phi) = \sup_N \mu(\Phi^N) &= \lim_N \mu(\Phi^N) \\
&= \lim_N \lim_n \mathrm{E}\, \Phi^N(., X_n) \\
&= \lim_n \lim_N \mathrm{E}\, \Phi^N(., X_n) \\
&= \lim_n \mathrm{E}\, \Phi(., X_n).
\end{aligned}$$

$\square$

Construction of the extended probability space: the processes Y, V and Z Recall that $V^{(n)} = \int_0^{\cdot} Z_s^{(n)}\, dW_s$. By Proposition 3.10, the sequence $(Y^{(n)}, V^{(n)})$, seen as a sequence of random variables with values in $\mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}}$, is tight. Let us denote $\mathbb{H} = \mathrm{L}_{\mathbb{L}}^2([0, T])$, and let $\mathbb{H}_\sigma$ be the space $\mathbb{H}$ endowed with its weak topology (note that this topology has the same Borel sets as

the strong topology). Each $Z^{(n)}$ can be considered as a random variable with values in $\mathbb{H}_\sigma$. Furthermore, by Corollary 3.6, the sequence $(Z^{(n)})$ is tight in $\mathbb{H}_\sigma$: Indeed, the closed balls are compact in $\mathbb{H}_\sigma$, and we have

$$\sup_n \mathrm{P}\left\{\left\|Z^{(n)}\right\|_{\mathbb{H}} \geq R\right\} \leq \sup_n \frac{1}{R^2}\,\mathrm{E}\int_0^T \left\|Z_s^{(n)}\right\|_{\mathrm{L}_{\mathbb{L}}^2[0,T]}^2\,ds$$
$$\rightarrow 0 \text{ when } R \rightarrow \infty.$$

Thus $(Y^{(n)}, V^{(n)}, Z^{(n)})$ is a tight sequence of $\mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma$-valued variables.

We now consider the space $\mathcal{Y}(\Omega, \mathcal{F}, \mathrm{P}; \mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma)$, which we denote for simplicity by $\mathcal{Y}$. By Prohorov's sequential compactness criterion for Young measures [12, Theorem 4.3.5], we can extract a subsequence of $(Y^{(n)}, V^{(n)}, Z^{(n)})$ (for simplicity, we denote this extracted sequence by $(Y^{(n)}, V^{(n)}, Z^{(n)})$) which converges stably to some $\mu \in \mathcal{Y}$, that is, for every measurable bounded mapping $\Phi : \Omega \times \mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma \mapsto \mathbb{R}$ such that $\Phi(\omega, ., ., .)$ is continuous for all ω, we have

$$(34) \quad \lim_{n \rightarrow \infty} \int_\Omega \Phi\left(\omega, Y^{(n)}(\omega), V^{(n)}(\omega), Z^{(n)}(\omega)\right)\,d\mathrm{P}(\omega)$$
$$= \int_\Omega \int_{\mathbb{D} \times \mathbb{D} \times \mathbb{H}} \Phi(\omega, y, v, z)\,d\mu_\omega(y, v, z)\,d\mathrm{P}(\omega).$$

In particular, $(Y^{(n)}, V^{(n)}, Z^{(n)})$ converges in law to the image of μ by the canonical projection of $\Omega \times \mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma$ on $\mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma$.

Let us denote by $\mathcal{D}$ the Borel σ-algebra of $\mathbb{D}$ (recall that S has the same Borel subsets as Skorokhod's J_1 topology), and, for each $t \in [0, T]$, let $\mathcal{D}_t$ be the sub-σ-algebra of $\mathcal{D}$ generated by the projection onto $\mathbb{D}_{\mathbb{R}^d}([0, t])$. Similarly, let $\mathcal{H}$ denote the Borel σ-algebra of $\mathbb{H}_\sigma$, and, for each $t \in [0, T]$, let $\mathcal{H}_t$ be the sub-σ-algebra of $\mathcal{H}$ generated by the projection onto $\mathrm{L}_{\mathbb{L}}^2([0, t])$. We define a stochastic basis $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$ by

$$\underline{\Omega} = \Omega \times \mathbb{D} \times \mathbb{D} \times \mathbb{H}, \quad \underline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{D} \otimes \mathcal{D} \otimes \mathcal{H}, \quad \underline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{D}_t \otimes \mathcal{D}_t \otimes \mathcal{H}_t,$$

and we define a process (Y, V, Z) on $\underline{\Omega}$ by

$$Y(\omega, y, v, z) = y, \quad V(\omega, y, v, z) = v, \quad Z(\omega, y, v, z) = z.$$

Clearly, (Y, V, Z) is $(\underline{\mathcal{F}}_t)$–adapted. Furthermore, the law of (Y, V, Z) is the marginal measure of μ on $\mathbb{D} \times \mathbb{D} \times \mathbb{H}$, in particular $(Y^{(n)}, V^{(n)}, Z^{(n)})$ converges in law to (Y, V, Z) on $\mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma$. By [21, Theorem 3.11], *we can (and will) furthermore choose the extracted sequence such that, there exists a countable*

33

set $\mathcal{N} \subset [0, T[$ such that, for every $t \in [0, T] \setminus \mathcal{N}$, the sequence $(Y_t^{(n)}, V_t^{(n)})$ converges in law to (Y_t, V_t).

Now, the random variables $(Y^{(n)}, V^{(n)}, Z^{(n)})$ can be seen as random elements defined on $\underline{\Omega}$, using the notations, for $n \geq 1$:

$$Y^{(n)}(\omega, y, v, z) := Y^{(n)}(\omega),$$
$$V^{(n)}(\omega, y, v, z) := V^{(n)}(\omega),$$
$$Z^{(n)}(\omega, y, v, z) := Z^{(n)}(\omega).$$

Furthermore, $(Y^{(n)}, V^{(n)}, Z^{(n)})$ is $(\underline{\mathcal{F}}_t)$–adapted for each n. Likewise, we set $W(\omega, y, v, z) = W(\omega)$.

Lemma 3.13 *The process W is an $(\underline{\mathcal{F}}_t)$–standard Brownian motion under the probability μ.*

Proof By Balder's result on K-convergence [5, 6], which is valid for Hausdorff spaces with metrizable compact subsets [12, Lemma 4.5.4], each subsequence of $(Y^{(n)}, V^{(n)})$ contains a further subsequence (Y^{n_k}, V^{n_k}) which K–converges to μ, that is, for each subsequence $(Y^{(n_k')}, V^{(n_k')})$ of $(Y^{(n_k)}, V^{(n_k)})$, we have

$$\lim_n \frac{1}{n} \sum_{k=1}^n \delta_{(Y^{(n_k')}(\omega), V^{(n_k')}(\omega))} = \mu_\omega \text{ a.e.}$$

where $\delta_{(y,v)}$ denotes the Dirac measure on (y, v) and the limit is taken in the narrow convergence. This entails that, for every $B \in \mathcal{D}_t \otimes \mathcal{D}_t \otimes \mathcal{H}_t$, the mapping $\omega \mapsto \mu_\omega(B)$ is $\mathcal{F}_t$–measurable. The result follows from Lemma 2.10. □

Properties of the processes Y and V

Lemma 3.14 *Let H and K be $\mathbb{R}^d$-valued random variables defined on $\underline{\Omega}$. Let $t \in [0, T]$. In order that H and K have the same conditional expectation with respect to $\underline{\mathcal{F}}_t$, it is sufficient that*

$$(35) \quad \int_\Omega \int_{\mathbb{D} \times \mathbb{D} \times \mathbb{H}} \Phi(\omega, y, v, z) H(\omega, y, v, z) \, d\mu_\omega(y, v, z) \, d\mathrm{P}(\omega)$$
$$= \int_\Omega \int_{\mathbb{D} \times \mathbb{D} \times \mathbb{H}} \Phi(\omega, y, v, z) K(\omega, y, v, z) \, d\mu_\omega(y, v, z) \, d\mathrm{P}(\omega)$$

for every bounded $\underline{\mathcal{F}}_t$-measurable function $\Phi : \underline{\Omega} \to \mathbb{R}$ such that $\Phi(\omega, ., ., .)$ is continuous for all $\omega \in \Omega$.

Proof Let $\mathcal{C}$ be the set of functions $\Phi : \underline{\Omega} \to \mathbb{R}$ which are $\underline{\mathcal{F}}_t$-measurable and such that $\Phi(\omega, ., ., .)$ is continuous for all $\omega \in \Omega$. The set $\mathcal{C}$ is stable by multiplication of two functions and generates $\underline{\mathcal{F}}_t$. Assume that (35) holds for every $\Phi \in \mathcal{C}$, and let $\mathcal{E}$ be the vector space of bounded $\underline{\mathcal{F}}_t$-measurable functions φ defined on $\underline{\Omega}$ such that

$$\int_\Omega \int_{\mathbb{D} \times \mathbb{D} \times \mathbb{H}} \varphi(\omega, y, v, z) H(\omega, y, v, z) \, d\mu_\omega(y, v, z) \, d\mathrm{P}(\omega)$$
$$= \int_\Omega \int_{\mathbb{D} \times \mathbb{D} \times \mathbb{H}} \varphi(\omega, y, v, z) K(\omega, y, v, z) \, d\mu_\omega(y, v, z) \, d\mathrm{P}(\omega).$$

The space $\mathcal{E}$ contains $\mathcal{C}$. Furthermore, $\mathcal{E}$ contains the constant functions and is stable under monotone limits of uniformly bounded sequences. By the monotone class theorem (see [34, Appendix A0] and [14, Theorème 21, page 20]), $\mathcal{E}$ contains all bounded $\underline{\mathcal{F}}_t$-measurable functions. $\square$

Lemma 3.15 *The process V is a martingale with respect to $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$.*

Proof Let $t \in [0, T]$, and let $s \in [0, T - t]$. By Lemma 3.14, in order to prove that $\mathrm{E}^{\underline{\mathcal{F}}_t}(V_{t+s}) = V_t$, we only need to show that, for each bounded $\underline{\mathcal{F}}_t$-measurable $\Phi : \underline{\Omega} \to \mathbb{R}$ such that $\Phi(\omega, ., ., .)$ is continuous for all $\omega \in \Omega$, we have

(36) $$\mathrm{E}\left(\Phi \times V_{t+s}\right) = \mathrm{E}\left(\Phi \times V_t\right).$$

Let us denote, for any $r \in [0, T]$, any $v \in \mathbb{D}$ and any $\delta > 0$

(37) $$\pi_r(v) = v(r) \quad \text{and} \quad \pi_{r,\delta}(v) = \frac{1}{\delta} \int_r^{r+\delta} v(s) \, ds.$$

The mapping $\pi_r : \mathbb{D} \to \mathbb{R}^d$ is not continuous for the topology S, but $\pi_{r,\delta}$ is S-continuous, and we have

$$\lim_{\delta \to 0} \pi_{r,\delta}(v) = \pi_r(v).$$

Let $\delta > 0$. Let

$$\phi(\omega, y, v, z) = \Phi(\omega, y, v, z) \left(\pi_{t+s,\delta}(v) - \pi_{t,\delta}(v)\right).$$

By Corollary 3.6, the sequence $(\phi(\omega, Y^{(n)}, V^{(n)}, Z^{(n)}))$ is bounded in $\mathrm{L}^2_{\mathbb{R}^d}(\Omega)$, thus it is uniformly integrable. We can thus apply Lemma 3.12 to the

integrand ϕ. Using the definition of V and the fact that each $V^{(n)}$ is a martingale, we get

$$
\mathrm{E}\left(\Phi \times \left(\frac{1}{\delta}\int_{t+s}^{t+s+\delta} V_u\,du - \frac{1}{\delta}\int_t^{t+\delta} V_u\,du\right)\right)
$$

$$
= \int_\Omega \int_{\mathbb{D}\times\mathbb{D}\times\mathbb{H}} \Phi(\omega,y,v,z)\,(\pi_{t+s,\delta} - \pi_{t,\delta})\,(v)\,d\mu_\omega\,(y,v,z)\;d\mathrm{P}(\omega)
$$

$$
= \lim_{n\to\infty}\int_\Omega \Phi\left(\omega, Y^{(n)}(\omega), V^{(n)}(\omega), Z^{(n)}(\omega)\right)\frac{1}{\delta}\int_t^{t+\delta}\left(V_{u+s}^{(n)}(\omega) - V_u^{(n)}(\omega)\right)\,du\,d\mathrm{P}(\omega)
$$

$$
= \lim_{n\to\infty}\int_\Omega \Phi\left(\omega, Y^{(n)}(\omega), V^{(n)}(\omega), Z^{(n)}(\omega)\right)\mathrm{E}^{\mathcal{F}_t}\left(\frac{1}{\delta}\int_t^{t+\delta}\left(V_{u+s}^{(n)} - V_u^{(n)}\right)\,du\right)\,d\mathrm{P}
$$

$$
= 0.
$$

We deduce that

$$
\mathrm{E}\left(\Phi \times (V_{t+s} - V_t)\right)
$$

$$
= \lim_{\delta\to 0}\mathrm{E}\left(\Phi \times \left(\frac{1}{\delta}\int_{t+s}^{t+s+\delta} V_u\,du - \frac{1}{\delta}\int_t^{t+\delta} V_u\,du\right)\right) = 0
$$

by boundedness in $\mathrm{L}^2_{\mathbb{R}^d}(\underline{\Omega,\mathcal{F}},\mu)$ of $(V_r)_{0\leq r\leqq T}$. $\qquad\square$

Lemma 3.16 *Let* $\widehat{V}_t = \int_0^t Z_s\,dW_s$. *The martingale* $L := V - \widehat{V}$ *is orthogonal to* W.

Proof Let us denote the coordinates processes as in the following examples:
$V = (V^{[i]})_{1\leq i\leq d}$, $Z_t^{(n)} = (Z_t^{(n),[i,k]})_{1\leq i\leq d, 1\leq k\leq m}$, $Z_t = (Z_t^{[i,k]})_{1\leq i\leq d, 1\leq k\leq m}$, $W_t = (W^{[k]})_{1\leq k\leq m}$.

Let $i \in \{1,\dots,m\}$ and $j \in \{1,\dots,d\}$. Let us denote by $[P,Q]$ the quadratic cross variation of two semimartingales P and Q. For each n, let

$$
N_t^{(n),[i,j]} = W_t^{[i]}V_t^{(n),[j]} - \left[W^{[i]}, V^{(n),[j]}\right]_t
$$

$$
= W_t^{[i]}\sum_{k=1}^m \int_0^t Z_r^{(n),[j,k]}\,dW_r^{[k]} - \int_0^t Z_r^{(n),[j,i]}\,dr
$$

$$
\widehat{N}_t^{[i,j]} = W_t^{[i]}\widehat{V}_t^{[j]} - \left[W^{[i]}, \widehat{V}^{[j]}\right]_t
$$

$$
= W_t^{[i]}\sum_{k=1}^m \int_0^t Z_r^{[j,k]}\,dW_r^{[k]} - \int_0^t Z_r^{[j,i]}\,dr.
$$

36

As W is continuous, the processes $N^{(n),[i,j]}$ and $\widehat{N}^{[i,j]}$ are continuous martingales. Let $\Phi : \underline{\Omega} \to \mathbb{R}$ be a bounded $\underline{\mathcal{F}}_t$-measurable function such that $\Phi(\omega,.,.,.)$ is continuous for all $\omega \in \Omega$. Observe that, from the stable convergence of $V^{(n)}$ to V, we have, for any $\tau \in [0,T]$ and any $\delta > 0$,

$$(38) \quad \lim_n \mathrm{E}\left(\pi_{\tau,\delta}(W^{[i]}V^{(n),[j]})\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right)$$
$$= \mathrm{E}\left(\pi_{\tau,\delta}(W^{[i]}V^{[j]})\Phi\right)$$

using Lemma 3.12 with the integrand $\phi(\omega,y,v,z) = \pi_{\tau,\delta}(W^{[i]}(\omega)v^{(n),[j]})\Phi(\omega,y,v,z)$, where $\pi_{\tau,\delta}$ is defined as in (37). Similarly, from the stable convergence of $(V^{(n)},Z^{(n)})$ to (V,Z), and applying Lemma 3.12 with the integrand

$$\phi(\omega,y,v,z) = \left(\int_0^\tau z_r^{[j,i]}\,dr\right)\Phi(\omega,y,v,z),$$

we get

$$(39) \quad \lim_n \mathrm{E}\left(\left(\int_0^\tau Z_r^{(n),[j,i]}\,dr\right)\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right)$$
$$= \mathrm{E}\left(\left(\int_0^\tau Z_r^{[j,i]}\,dr\right)\Phi\right).$$

Let $t \in [0,T]$, and let $s \in [0,T-t]$. Using (38), (39), and the fact that

$N^{(n),[i,j]}$ and $\widehat{N}^{[i,j]}$ are martingales, we get, for any $\delta > 0$,

$$\frac{1}{\delta} \int_t^{t+\delta} \mathrm{E}\left(W_{u+s}^{[i]}\left(V_{u+s}^{[j]} - \widehat{V}_{u+s}^{[j]}\right)\Phi\right) du$$

$$= \frac{1}{\delta} \int_t^{t+\delta} \lim_n \left[\mathrm{E}\left(W_{u+s}^{[i]} V_{u+s}^{(n),[j]}\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right) \right.$$

$$\left. - \mathrm{E}\left(W_{u+s}^{[i]} \widehat{V}_{u+s}^{[j]}\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right) \right] du$$

$$= \frac{1}{\delta} \int_t^{t+\delta} \lim_n \left[\mathrm{E}\left(\left(N_{u+s}^{(n),[i,j]} + \int_0^{u+s} Z_r^{(n),[j,i]}\,dr\right)\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right) \right.$$

$$\left. - \mathrm{E}\left(\left(\widehat{N}_{u+s}^{[i,j]} + \int_0^{u+s} Z_r^{[j,i]}\,dr\right)\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right) \right] du$$

$$= \frac{1}{\delta} \int_t^{t+\delta} \lim_n \mathrm{E}\left(\left(N_{u+s}^{(n),[i,j]} - \widehat{N}_{u+s}^{[i,j]}\right)\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right) du$$

$$= \frac{1}{\delta} \int_t^{t+\delta} \lim_n \mathrm{E}\left(\left(N_u^{(n),[i,j]} - \widehat{N}_u^{[i,j]}\right)\Phi\left(.,Y^{(n)},V^{(n)},Z^{(n)}\right)\right) du$$

$$= \frac{1}{\delta} \int_t^{t+\delta} \mathrm{E}\left(W_u^{[i]}\left(V_u^{[j]} - \widehat{V}_u^{[j]}\right)\Phi\right) du.$$

Passing to the limit when $\delta \to 0$ yields

$$\mathrm{E}\left(W_{t+s}^{[i]}\left(V_{t+s}^{[j]} - \widehat{V}_{t+s}^{[j]}\right)\Phi\right) = \mathrm{E}\left(W_t^{[i]}\left(V_t^{[j]} - \widehat{V}_t^{[j]}\right)\Phi\right).$$

By Lemma 3.14, this shows that $W^{[i]}\left(V^{[j]} - \widehat{V}^{[j]}\right)$ is a martingale.

$\square$

3.5 Proof of the main result

In this part, we use the special form of f with respect to Z: By hypothesis (H_2), f has the form

(40) $$f(s,x,y,z) = \alpha(s,x,y)z + \beta(s,x,y),$$

where α and β are bounded and continuous in (x,y), and α takes its values in the space $\mathrm{L}(\mathbb{L},\mathbb{R}^d)$ of linear mappings from $\mathbb{L}$ to $\mathbb{R}^d$.

We first prove a technical lemma.

Lemma 3.17 *Let $\mathbb{K}$ be the space of linear mappings from $\mathbb{R}^d$ to $\mathbb{R}^l$ for some $l \geq 1$. Let $b : [0, T] \to \mathbb{K}$ be a continuous function. For each $t \in [0, T]$, the mapping*

$$\Phi : \begin{cases} \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma & \to & \mathbb{R}^l \\ (y, z) & \mapsto & \int_0^t b(s).f(s, x(s), y(s), z(s)) \, ds \end{cases}$$

is sequentially continuous. Furthermore, if $y_n \to y$ in $\mathbb{D}_{\mathrm{S}}$ and $z_n \to z$ in $\mathbb{H}_\sigma$, then, for every $t \in [0, T]$, we have

$$\lim_n \left(\int_0^t b(s).f(s, x(s), y_n(s), z_n(s + 1/n)) \, ds \right.$$
$$\left. - \int_0^t b(s).f(s, x(s), y_n(s), z_n(s)) \, ds \right) = 0.$$

Proof We only need to prove the lemma for $f(s, x, y, z) = \alpha(s, x, y)z$. As x does not play any role in our reasoning, we write for simplicity $f(s, x, y, z) = \alpha(s, y)z$.

First, for every $z \in \mathrm{L}^2_{\mathbb{L}}[0, T]$, we have

(41) $$\lim_{n \to \infty} \|z - z(. + 1/n)\|_{\mathrm{L}^2_{\mathbb{L}}[0,T]} = 0.$$

Indeed, for every $\epsilon > 0$, there exists a continuous function $u : [0, T] \to \mathbb{L}$ such that $\|z - u\|_{\mathrm{L}^2_{\mathbb{L}}[0,T]} < \epsilon$. Then we have, for every $n \geq 1$,

$$\|z(. + 1/n) - u(. + 1/n)\|_{\mathrm{L}^2_{\mathbb{L}}[0,T]} < \epsilon.$$

But the family $u(. + 1/n)$ is uniformly integrable because it is bounded in $\mathrm{L}^2_{\mathbb{L}}[0, T]$, thus, by Vitali's theorem and the continuity of u,

$$\lim_{n \to \infty} \|u - u(. + 1/n)\|_{\mathrm{L}^2_{\mathbb{L}}[0,T]} = 0.$$

We conclude by the triangular inequality that

$$\limsup_{n \to \infty} \|z - z(. + 1/n)\|_{\mathrm{L}^2_{\mathbb{L}}[0,T]} \leq 2\epsilon,$$

which proves (41). This entails that, for every $b \in \mathbb{D}_{\mathbb{R}^d}[0, T]$, and for all $y \in \mathbb{D}_{\mathbb{R}^d}[0, T]$, $z \in \mathrm{L}^2_{\mathbb{L}}[0, T]$, and $t \in [0, T]$,

(42) $$\lim_{n \to \infty} \int_t^T b(s).(\alpha(s, y(s))z(s + 1/n) - \alpha(s, y(s))z(s)) \, ds = 0$$

39

because b and α are bounded.

Now, let $y_n \to y$ in $\mathbb{D}_S$ and $z_n \to z$ in $\mathbb{H}_\sigma$. We have in particular

$$y_n(s) \to y(s) \text{ for a.e. } s \in [0, T] \text{ and } \sup_n \|z_n\|_{\mathbb{H}} < +\infty,$$

thus

$$\|\Phi(y_n, z_n)(t) - \Phi(y, z)(t)\|$$
$$= \left\| \int_0^t b(s).\left(\alpha(s, y_n(s)) - \alpha(s, y(s))\right) z_n(s)\, ds \right.$$
$$\left. + \int_0^t b(s).\alpha(s, y(s))\left(z_n(s) - z(s)\right) ds \right\|$$
$$\leq \sup_n \|z_n\|_{\mathbb{H}} \left(\int_0^t \|b(s)\| \|\alpha(s, y_n(s)) - \alpha(s, y(s))\|^2\, ds \right)^{1/2}$$
$$+ \left\| \int_0^t b(s).\alpha(s, y(s))\left(z_n(s) - z(s)\right) ds \right\|$$
$$\to 0 \text{ when } n \to \infty.$$

A similar calculation yields

$$\lim_n \left\| \int_0^t b(s).\alpha(s, y_n(s))\left(z_n(s + 1/n) - z_n(s)\right) ds \right\|$$
$$= \lim_n \left\| \int_0^t b(s).\left(\alpha(s, y_n(s)) - \alpha(s, y(s))\right)\left(z_n(s + 1/n) - z_n(s)\right) ds \right.$$
$$\left. + \int_0^t b(s).\alpha(s, y(s))\left(z_n(s + 1/n) - z_n(s)\right) ds \right\|$$
$$\leq 2 \sup_n \|z_n\|_{\mathbb{H}} \left(\int_0^t \|b(s)\| \|\alpha(s, y_n(s)) - \alpha(s, y(s))\|^2\, ds \right)^{1/2}$$
$$+ \left\| \int_0^t b(s).\alpha(s, y(s))\left(z_n(s + 1/n) - z_n(s)\right) ds \right\|$$
$$\to 0 \text{ when } n \to \infty$$

by the dominated convergence theorem and (42). $\square$

In order to check that (Y, Z) is a solution to (2), we prove in the next lemma that we can replace $\widetilde{Z}^{(n)}$ by $Z^{(n)}$ in the limit of $\int_t^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)})\, ds$.

40

Lemma 3.18 *For each $t \in [0, T]$, the sequence*

$$\int_t^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds - \int_t^T f(s, X_s, Y_s^{(n)}, Z_s^{(n)}) \, ds$$

converges to 0 in μ-probability.

Proof With the notations of (40), we only need to check that

$$\int_t^T \alpha(s, X_s, Y_s^{(n)})\big(\widetilde{Z}_s^{(n)} - Z_s^{(n)}\big) \, ds$$

converges to 0 in probability.

Now, by Lemma 3.4 and Proposition 3.5, the sequence $(\widetilde{Z}^{(n)})$ is bounded in $\mathrm{L}_{\mathbb{L}}^2(\Omega \times [0, T])$, thus it can be viewed as a tight sequence of $\mathbb{H}_\sigma$-valued random variables. Enlarging the space $\underline{\Omega}$ to $\Omega \times \mathbb{D} \times \mathbb{D} \times \mathbb{H} \times \mathbb{H}$, we can assume that $(Y^{(n)}, V^{(n)}, Z^{(n)}, \widetilde{Z}^{(n)})$ converges to a Young measure, still denoted by μ, in $\mathcal{Y}(\Omega, \mathcal{F}, \mathrm{P}; \mathbb{D}_\mathrm{S} \times \mathbb{D}_\mathrm{S} \times \mathbb{H}_\sigma \times \mathbb{H}_\sigma)$. We set

$$\widetilde{Z}(\omega, y, v, z, \widetilde{z}) = \widetilde{z}$$

and we extend Y, V, Z, and the σ-algebra $\underline{\mathcal{F}}_t$ in the obvious way.

Let K be an $\underline{\mathcal{F}}_t$-adapted process with càdlàg trajectories in $\mathbb{R}^d$, and assume that K is continuous with respect to y, v, z, and $\widetilde{z}$, and that the sequence

$$\Big(\int_0^T K_s.\alpha(s, X_s, Y_s)(Z_s^{(n)} - \widetilde{Z}_s^{(n)}) \, ds\Big)$$

41

is uniformly integrable. We have

$$\mathrm{E}\int_0^T K_s.\alpha(s, X_s, Y_s)(Z_s - \widetilde{Z}_s)\, ds$$

$$= \lim_n \mathrm{E}\int_0^T K_s.\alpha(s, X_s, Y_s^{(n)})(Z_s^{(n)} - \widetilde{Z}_s^{(n)})\, ds$$

(by Lemma 3.17 and Lemma 3.12, with

$$\Phi(\omega, y, v, z, \widetilde{z}) = \int_0^T K_s(\omega, y, v, z, \widetilde{z}).\alpha(s, x_s, y_s)(z_s - \widetilde{z}_s)\, ds))$$

$$= \lim_n \mathrm{E}\int_0^T K_s.\alpha(s, X_s, Y_s^{(n)})(Z_{s+1/n}^{(n)} - \widetilde{Z}_s^{(n)})\, ds$$

(by Lemma 3.17)

$$= \lim_n \mathrm{E}\int_0^T \mathrm{E}^{\mathcal{F}_s}\left(K_s.\alpha(s, X_s, Y_s^{(n)})(Z_{s+1/n}^{(n)} - \widetilde{Z}_s^{(n)})\right)\, ds$$

$$= \lim_n \mathrm{E}\int_0^T K_s.\alpha(s, X_s, Y_s^{(n)})(\widetilde{Z}_s^{(n)} - \widetilde{Z}_s^{(n)})\, ds$$

$$= 0.$$

In particular, one can take

$$K_s = \frac{\alpha(s, X_s, Y_s)(Z_s - \widetilde{Z}_s)}{1 + \left(\alpha(s, X_s, Y_s)(Z_s - \widetilde{Z}_s)\right)^2}.$$

Thus $\alpha(s, X_s, Y_s)(Z_s - \widetilde{Z}_s) = 0$, μ-a.e., for almost every $s \in [0, T]$.

Let $\Psi : \mathbb{R}^d \to \mathbb{R}$ be a bounded continuous function. Let

$$\Phi : \left\{ \begin{array}{ccc} \mathbb{D}_S \times \mathbb{D}_S \times \mathbb{H}_\sigma \times \mathbb{H}_\sigma & \to & \mathbb{R} \\ (x, y, z, \widetilde{z}) & \mapsto & \Psi\left(\int_t^T \alpha(s, x_s, y_s)(z_s - \widetilde{z}_s)\, ds\right). \end{array} \right.$$

By Hypothesis (H_2) and Lemma 3.17, Φ is sequentially continuous, thus, from the $\mathcal{F}$-stable convergence of $(X, Y^{(n)}, Z^{(n)}, \widetilde{Z}^{(n)})$,

$$\Psi(0) = \mathrm{E}\,\Psi\left(\int_t^T \alpha(s, X_s, Y_s)(Z_s - \widetilde{Z}_s)\, ds\right) = \lim_n \mathrm{E}\,\Phi(X, Y^{(n)}, Z^{(n)}, \widetilde{Z}^{(n)})$$

$$= \lim_n \mathrm{E}\,\Psi\left(\int_t^T \alpha(s, X_s, Y_s^{(n)})(\widetilde{Z}_s^{(n)} - Z_s^{(n)})\, ds\right).$$

This shows that the sequence

$$\left(\int_t^T \alpha(s, X_s, Y_s^{(n)}) \big(\widetilde{Z}_s^{(n)} - Z_s^{(n)} \big) \, ds \right)$$

converges to 0 in law, thus in probability. $\qquad\square$

Lemma 3.19 *The sequence* $(\int_{\cdot}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds)$ *converges in law to* $\int_{\cdot}^T f(s, X_s, Y_s, Z_s) \, ds$.

Proof By Lemma 3.9, we know that the sequence $(\int_{\cdot}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds)$ is relatively compact in law, thus we only need to show that it has only one possible limit in law, and that this limit is the law of $\int_{\cdot}^T f(s, X_s, Y_s, Z_s) \, ds$. By Lemma 3.18, it suffices to prove that $(\int_t^T f(s, X_s, Y_s^{(n)}, Z_s^{(n)}) \, ds)$ converges in law to $\int_t^T f(s, X_s, Y_s, Z_s) \, ds$ for each $t \in [0, T]$.

Let $\Psi : \mathbb{R}^d \to \mathbb{R}$ be a bounded continuous function. Let

$$\Phi : \begin{cases} \mathbb{D}_{\mathrm{S}} \times \mathbb{D}_{\mathrm{S}} \times \mathbb{H}_\sigma & \to \quad \mathbb{R} \\ (x, y, z) & \mapsto \quad \Psi \left(\int_t^T f(s, x_s, y_s, z_s) \, ds \right). \end{cases}$$

By Hypothesis (H_2) and Lemma 3.17, Φ is sequentially continuous, thus, from the $\mathcal{F}$-stable convergence of $(X, Y^{(n)}, Z^{(n)})$,

$$\mathrm{E}\, \Psi \left(\int_t^T f(s, X_s, Y_s, Z_s) \, ds \right) = \mu(\Phi) = \lim_n \mathrm{E}\, \Phi(X, Y^{(n)}, Z^{(n)})$$

$$= \lim_n \mathrm{E}\, \Psi \left(\int_t^T f(s, X_s, Y_s^{(n)}, Z_s^{(n)}) \, ds \right).$$

$\qquad\square$

Proof of Theorem 3.1 By Lemma 3.13, W is a Brownian motion on $(\Omega, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$. Let $L_t = V_t - V_0 - \widehat{V}_t$, $0 \le t \le T$. We have $L_0 = 0$ and L is a càdlàg martingale by Lemma 3.15, furthermore L is orthogonal to W by Lemma 3.16. Thus there only remains to prove that (Y, Z, L) satisfies (1).

Thanks to Proposition 3.10 and Lemma 3.9, we know that the sequence

$$(43) \qquad \left(X, Y^{(n)}, \int_{\cdot}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds, \int_{\cdot}^T Z_s^{(n)} \, dW_s \right)_{n \ge 1}$$

is tight in $C_\mathbb{M}[0, T] \times \mathbb{D}_S \times C_{\mathbb{R}^d}[0, T] \times \mathbb{D}_S$. Furthermore, $\left(\int_{\cdot}^T Z_s^{(n)} \, dW_s \right)_{n \geq 1}$ converges in law to $V_T - V_\cdot$, and, by Lemma 3.19, $\left(\int_{\cdot}^T f(s, X_s, Y_s^{(n)}, \widetilde{Z}_s^{(n)}) \, ds \right)_{n \geq 1}$ converges in law to $\int_{\cdot}^T f(s, X_s, Y_s, Z_s) \, ds$. Extracting if necessary a further subsequence, we can thus assume that the sequence (43) jointly converges in law on $C_\mathbb{M}[0, T] \times \mathbb{D}_S \times C_{\mathbb{R}^d}[0, T] \times \mathbb{D}_S$ to

$$\left(X, Y, \int_{\cdot}^T f(s, X_s, Y_s, Z_s) \, ds, V_T - V_\cdot \right).$$

Then the process

$$U_\cdot^{(n)} = Y_\cdot^{(n)} - \xi - \int_{\cdot}^T f(s, X_s, Y_s^{(n)}, Z_s^{(n)}) \, ds + \int_{\cdot}^T Z_s^{(n)} \, dW_s$$

converges in law in $\mathbb{D}_S$ to

$$\begin{aligned}
U_\cdot :=\, & Y_\cdot - \xi - \int_{\cdot}^T f(s, X_s, Y_s, Z_s) \, ds + V_T - V_\cdot \\
=\, & Y_\cdot - \xi - \int_{\cdot}^T f(s, X_s, Y_s, Z_s) \, ds + \int_{\cdot}^T Z_s \, dW_s + L_T - L_\cdot
\end{aligned}$$

But, by Lemma 3.7, $(\sup_{0 \leq t \leq T} U_t^{(n)})$ converges to 0 in probability, thus $U = 0$ a.e., which proves Theorem 3.1. $\qquad\square$

References

[1] David J. Aldous. Stopping times and tightness. *Ann. Probab.*, 6:335–340, 1978.

[2] Fabio Antonelli and Jin Ma. Weak solutions of forward-backward SDE's. *Stochastic Anal. Appl.*, 21(3):493–514, 2003.

[3] K. Bahlali, B. Mezerdi, M. N'zi, and Y. Ouknine. Weak solutions and a Yamada-Watanabe theorem for FBSDEs. *Random Oper. Stoch. Equ.*, 15(3):271–285, 2007.

[4] Khaled Bahlali, Boulekhrass Gherbal, and Brahim Mezerdi. Existence of optimal controls for systems driven by FBSDEs. *Systems Control Lett.*, 60:344–349, 2011.

[5] E. J. Balder. On Prohorov's theorem for transition probabilities. *Sém. Anal. Convexe*, 19:9.1–9.11, 1989.

[6] E. J. Balder. New sequential compactness results for spaces of scalarly integrable functions. *J. Math. Anal. Appl.*, 151:1–16, 1990.

[7] E. J. Balder. Lectures on Young measure theory and its applications in economics. *Rend. Istit. Mat. Univ. Trieste*, 31, suppl.:1–69, 2000. Workshop di Teoria della Misura et Analisi Reale Grado, 1997 (Italia).

[8] J. K. Brooks and N. Dinculeanu. Regularity and the Doob-Meyer decomposition of abstract quasimartingales. In *Seminar on Stochastic Processes, 1987 (Princeton, NJ, 1987)*, volume 15 of *Progr. Probab. Statist.*, pages 21–63. Birkhäuser Boston, Boston, MA, 1988.

[9] R. Buckdahn and H.-J. Engelbert. A backward stochastic differential equation without strong solution. *Teor. Veroyatn. Primen.*, 50(2):390–396, 2005.

[10] R. Buckdahn and H.-J. Engelbert. On the continuity of weak solutions of backward stochastic differential equations. *Teor. Veroyatn. Primen.*, 52(1):190–199, 2007.

[11] R. Buckdahn, H.-J. Engelbert, and A. Răşcanu. On weak solutions of backward stochastic differential equations. *Teor. Veroyatn. Primen.*, 49(1):70–108, 2004.

[12] Charles Castaing, Paul Raynaud de Fitte, and Michel Valadier. *Young measures on Topological Spaces. With Applications in Control Theory and Probability Theory*, volume 571 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2004.

[13] F. Delarue and G. Guatteri. Weak existence and uniqueness for forward-backward SDEs. *Stochastic Process. Appl.*, 116(12):1712–1742, 2006.

[14] Claude Dellacherie and Paul André Meyer. *Probabilités et Potentiel. Chapitres I à IV*. Hermann, Paris, 1975.

[15] J. L. Doob. *Stochastic processes*. John Wiley & Sons Inc., New York, 1953.

[16] H. J. Engelbert. On the theorem of T. Yamada and S. Watanabe. *Stochastics Stochastics Rep.*, 36(3-4):205–216, 1991.

[17] Jean Jacod. Weak and strong solutions of stochastic differential equations. *Stochastics*, 3(3):171–191, 1980.

[18] Jean Jacod. Théorèmes limite pour les processus. In P. L. Hennequin, editor, *École d'Été de Probabilités de Saint–Flour XIII-1983*, number 1117 in Lecture Notes in Math., pages 298–409, Berlin, 1985. Springer Verlag.

[19] Jean Jacod and Jean Mémin. Weak and strong solutions of stochastic differential equations: existence and stability. In *Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980)*, volume 851 of *Lecture Notes in Math.*, pages 169–212. Springer, Berlin, 1981.

[20] A. Jakubowski. The almost sure Skorokhod representation for subsequences in nonmetric spaces. *Theory Proba. Appl.*, 42(1):209–216, 1997.

[21] A. Jakubowski. A non–Skorohod topology on the Skorohod space. *Electron. J. Probab.*, 2, 1997. Paper 4, 21 pages.

[22] A. Jakubowski, J. Mémin, and G. Pagès. Convergence en loi des suites d'intégrales stochastiques sur l'espace $\mathbb{D}^1$ de Skorokhod. *Probab. Theory Relat. Fields*, 81(1):111–137, 1989.

[23] Adam Jakubowski, Mikhail I. Kamenskiĭ, and Paul Raynaud de Fitte. Existence of weak solutions to stochastic differential inclusions. *Stoch. Anal. Appl.*, 23(4):723–749, 2005.

[24] John L. Kelley. *General Topology.* Springer Verlag, Berlin, 1955.

[25] Thomas G. Kurtz. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.*, 12:951–965 (electronic), 2007.

[26] J. P. Lepeltier and J. San Martín. Backward stochastic differential equations with continuous coefficient. *Statist. Probab. Lett.*, 32(4):425–430, 1997.

[27] Jin Ma, Jianfeng Zhang, and Ziyu Zheng. Weak solutions for forward-backward SDEs—a martingale problem approach. *Ann. Probab.*, 36(6):2092–2125, 2008.

[28] Jean Mémin and Leszek Słomiński. Condition UT et stabilité en loi des solutions d'équations différentielles stochastiques. In *Séminaire de Probabilités, XXV*, volume 1485 of *Lecture Notes in Math.*, pages 162–177. Springer, Berlin, 1991.

[29] P.-A. Meyer and W. A. Zheng. Tightness criteria for laws of semi-martingales. *Ann. Inst. H. Poincaré Probab. Statist.*, 20(4):353–372, 1984.

[30] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.

[31] J. Pellaumail. Weak solutions for semimartingales. *Canad. J. Math.*, 33(5):1165–1181, 1981.

[32] Alfred Rényi. On stable sequences of events. *Sankhyā Ser. A*, 25:293–302, 1963.

[33] Laurent Schwartz. *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. Tata Institute of Fundamental Research Studies in Mathematics. Oxford University Press, London, 1973.

[34] Michael Sharpe. *General Theory of Markov Processes*. Number 133 in Pure and Applied Math. Academic Press, Boston, MA, 1988.

[35] Christophe Stricker. Lois de semimartingales et critères de compacité. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 209–217. Springer, Berlin, 1985.

[36] Michel Valadier. Désintégration d'une mesure sur un produit. *C. R. Acad. Sci. Paris Sér. I*, 276:A33–A35, 1973.

[37] Michel Valadier. A course on Young measures. *Rendiconti dell'istituto di matematica dell'Università di Trieste*, 26, suppl.:349–394, 1994. Workshop di Teoria della Misura et Analisi Reale Grado, 1993 (Italia).